INDUSTRIAL ELECTRONICS
AND CONTROL

PRINCIPLES OF ELECTRONICS.

BASIC ELECTRONICS. In preparation.

No. 879,532. PATENTED FEB. 18, 1908

L. DE FOREST.

SPACE TELEGRAPHY.

APPLICATION FILED JAN 29 1907

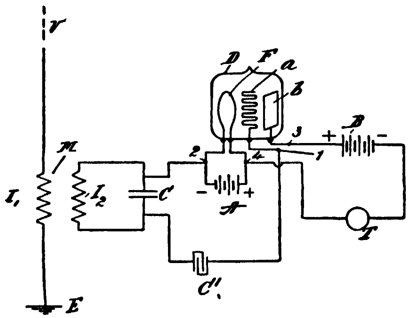
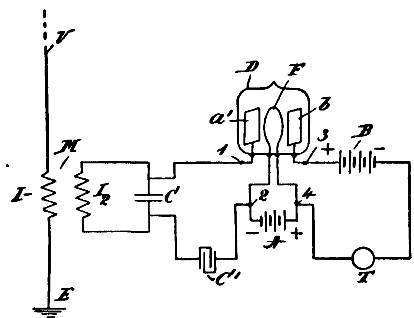


Fig.1.



WITNESSES- Fig.2. INVENTOR:

E. B. Tomlinson

Patrick J. Conroy

Lee de Forest

by Geo. K. Woodworth,

Atty.

Frontispiece

INDUSTRIAL ELECTRONICS AND CONTROL

ROYCE GERALD KLOEFFLER
Head, Department of Electrical Engineering
Kansas State College

JOHN WILEY & SONS, INC., NEW YORK
CHAPMAN & HALL, LIMITED, LONDON

PRINTED IN THE UNITED STATES OF AMERICA

To my wife Hazel Marguerite for her patience and inspiration while the manuscript was in preparation, and to my daughter Doris May who made many of the drawings for this book.

PREFACE

This book has been written to give the college engineering student a survey of the theory and applications of electronics in industry. Most of the books on electronics published for college students have been prepared for the communications major in electrical engineering. This textbook has been designed to meet the requirements of the electrical power major and of students in mechanical and chemical engineering who desire a knowledge of industrial electronic applications. No attempt has been made to provide a quantitative approach to the design of circuits since such calculations require several credit-hours of basic studies in the field. The early chapters of the text are for the benefit of the student whose previous training has not included the basic theory of electron tubes. For students who have had such training the first five chapters may be omitted. Some teachers may find the first eight chapters suitable for an introductory course in electron tubes.

I have employed the ASA standard symbols for all circuit diagrams. The electron current flow has been used in the explanation of the operation of electron tubes since it appears more logical. In circuit diagrams covering electronic power applications polarities have been carefully marked and the direction of both electron and conventional current are frequently indicated. Simple basic circuits have been employed wherever possible rather than specialized ones.

I am grateful to my colleagues, Professors Karl H. Martin and J. Edmond Wolfe, for their helpful criticisms and suggestions, and to many engineering friends in industry who have read portions of the text and have given permission for quotation from their articles or books.

R. G. Kloeffler

January 1949

CONTENTS

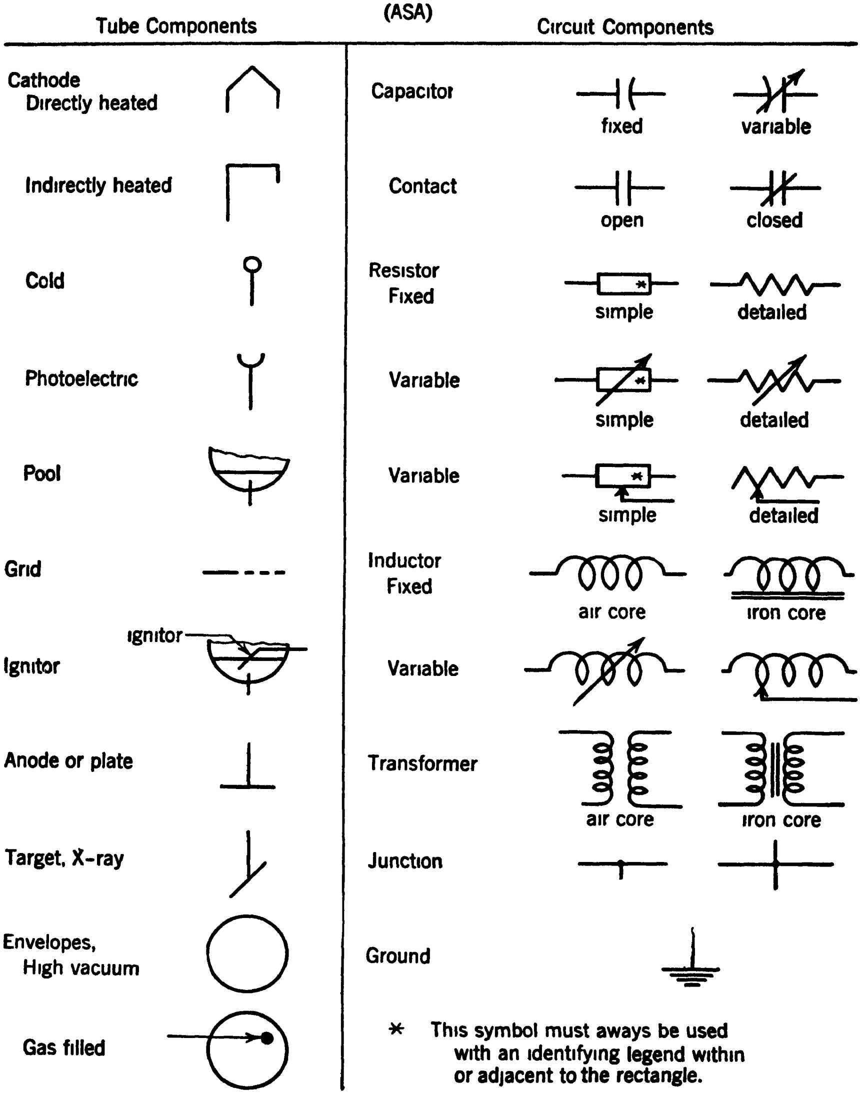
CHAPTER		PAGE
	STANDARD SYMBOLS FOR ELECTRON-TUBE CIRCUITS	xiii
	ELECTRICAL GRAPHICAL SYMBOLS (ASA)	xiv
I	INTRODUCTION	1
II	ELECTRON EMISSION	19
III	VACUUM DIODES	39
IV	GRID-CONTROLLED VACUUM TUBES	55
V	ELECTRICAL CONDUCTION IN GASES	95
VI	GASEOUS AND VAPOR ELECTRON TUBES	110
VII	PHOTOELECTRICITY	158
VIII	CRYSTAL AND METALLIC RECTIFIERS	180
IX	ALTERNATING-CURRENT THEORY	196
X	COMPONENTS AND CIRCUITS FOR CONTROL	205
XI	PRINCIPLES OF CONTROL AND SERVOMECHANISMS	249
XII	RECTIFICATION AND INVERSION	268
XIII	HIGH-FREQUENCY HEATING	297
XIV	RESISTANCE WELDING	328
XV	ELECTROSTATIC PRECIPITATION	362
XVI	ELECTRONIC OPERATION OF DIRECT-CURRENT MOTORS	372
XVII	PHOTOELECTRIC CONTROL DEVICES	387
XVIII	ELECTRONIC REGULATORS	407
XIX	ELECTRONIC POWER CONTROLS	415
XX	AMPLIDYNE SERVOMECHANISMS	422
XXI	X-RAY APPLICATIONS	430
XXII	SPECIAL PHOTO APPLICATIONS	451
	VACUUM ELECTRONIC TUBES (Table)	472
	GAS-FILLED ELECTRONIC TUBES (Table)	473
	INDEX	475

STANDARD SYMBOLS FOR ELECTRON-TUBE CIRCUITS

Component	Grid Voltage	Plate Voltage	Grid Current	Plate Current	Load Voltage
Total value, instantaneous	e_c	e_b	i_c	i_b	e_L
Total value, instantaneous maximum	E_{cm}	E_{bm}	I_{cm}	I_{bm}	E_{Lm}
Total value, average	E_c	E_b	I_c	I_b	E_L
Quiescent or zero signal average value	E_{co}	E_{bo}	I_{co}	I_{bo}	E_{Lo}
Varying component, instantaneous	e_g	e_p	i_g	i_p	e_z
Varying component, instantaneous maximum	E_{gm}	E_{pm}	I_{gm}	I_{pm}	E_{zm}
Varying component, effective	E_g	E_p	I_g	I_p	E_z
Varying component, average	E_{go}	E_{po}	I_{go}	I_{po}	E_{zo}

Supply voltages:	
Grid (d-c)	E_{cc} or E_{cc1}
Screen grid (d-c)	E_{cc2}
Plate (d-c)	E_{bb}
Filament or heater	E_{ff}
Filament or heater terminal voltage	E_f
Filament or heater current	I_f

ELECTRICAL GRAPHICAL SYMBOLS
(ASA)



Chapter I

INTRODUCTION

Electronics is a term signifying certain developments that have centered around several discoveries and inventions made near the close of the nineteenth and the beginning of the twentieth centuries. In 1887 Hertz discovered the Hertzian waves. In 1895 Roentgen invented the X-ray tube. About 1898 Marconi demonstrated the possibilities of wireless communication. In 1902 Fleming invented the "valve" or two-electrode detector. In 1906 Deforest invented the audion or three-electrode tube. These basic discoveries gave an impetus to the work of hundreds of other scientists whose cumulative inventions and developments have produced what is known as electronics. The layman may consider electronics as a combination of X rays, radio, telephone repeaters, electric eye, television, and radar. This statement fails to include the important applications in the industrial field, such as high-frequency heating, power rectification, resistance welding, and electronic control. This textbook will treat of those applications of electronics and control that lie outside the field of communication. These industrial applications were developed later than those in the communication field, yet they are proving of equal importance in the material progress of our present-day civilization. Communication brings the men of the world closer together, whereas electronic and other new controls lighten the work of man and improve the quality of his handiwork.

Electronics is that branch of science and technology which relates to the conduction of electricity through gases and in vacuo.* This definition involves the flow of electrons in vacuum tubes and the movement of electrons and ions in tubes containing gas or vapor at low pressures. The term also covers the action taking place in all circuits associated with these electron tubes. Hence, in a broader sense, electronics may be considered to include nearly all electrical phenomena.

The movement of electrons and ions in electron tubes involves cer-

* Definition approved by the American Standards Association.

tain physical phenomena not considered in the study of electrical machinery. These phenomena include (1) the removal of electrons from solids, (2) the production of ions in gases, (3) the movement of electrons and ions in space between electrodes, and (4) the control of the flow of electrons and ions by electrostatic and magnetic fields. An understanding of these processes may be aided by a review of some chemical, physical, and electrical concepts.

Atomic Structure. The electron theory of electricity and matter is a product of our twentieth-century thinking and research. At the beginning of the century Thompson suggested the electron as a part of the normal atom. In 1913 Robert Millikan published the result of his work on isolating and measuring the charge on the ion. Thus the electron, a fundamental indivisible particle carrying a negative charge, was discovered. The counterpart or mate of the electron was named the proton and consists of a particle having a mass approximately 1840 times that of the electron and a positive charge equal in magnitude to the charge on the electron. With these two particles as the building blocks of nature, Bohr suggested the structure of the atom as shown in Fig. 1. He pictured that atom as consisting of a small dense core or nucleus about which one or more electrons revolve. This structure is analogous to our solar system; the nucleus corresponds to the sun, and the revolving electrons correspond to the earth and the planets. The nucleus of the atom contains all the protons and usually some of the electrons for the particular element involved. For the simple hydrogen atom the single proton constitutes the nucleus and the single electron the lone planet. For helium the nucleus consists of four protons and two electrons; the other two electrons revolve around the nucleus and serve as the planets of the system. The attraction of the positive nucleus for the revolving electrons is counterbalanced by the centrifugal force of their motion about the nucleus. Bohr assumed the paths or orbits of the electrons to be circles and ellipses. The electrons moving in orbits close to the nucleus have large forces acting upon them, whereas those in outer orbits are acted upon by much smaller forces. The amount of energy possessed by an electron revolving in any orbit is definite and characteristic of that orbit. To explain the

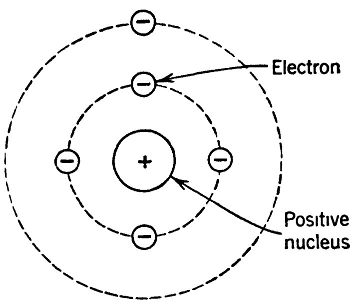


Fig. 1. Bohr's planetary structure of an atom.

property of radiation due to electrons, it was necessary to assume that an electron may have several orbits and that it is capable of jumping from one to another of these orbits under suitable excitation. The change from one orbit to another is accompanied by the absorption or radiation of energy. This radiated energy may be in the form of light, heat, or other wave energy. Because of this energy exchange some scientists have preferred to speak of the different orbits as energy levels.

The nucelus and electrons within an atom are very small and very far apart. The radius of the simplest atomic structure, the hydrogen atom, is 10^{-8} centimeter, and the radius of its orbital electron is 2×10^{-13} centimeter. The radius of the nucleus (lone proton) will be $1/_{1840}$ that of the electron. A conception of the relative magnitude of this hydrogen atomic system may be obtained by expanding the nucleus to the size of a baseball located at the geographical center of the United States (near Manhattan, Kansas). The revolving electron will pass through New York City and San Francisco and will be a sphere 300 feet in diameter—big enough to fill an average-size stadium or baseball park. Thus it is apparent that an atom is a hollow nebulous sphere—a swarm of specks occupying a small part of space. This conception is very helpful in understanding ionization in gases and many phenomena of electronics.

Two decades after Bohr's picture of atomic structure was offered, Carl Anderson discovered the positron, a positive particle having the same mass and magnitude of charge as the electron. About the same time several scientists isolated the neutron, a particle having the same mass as the proton but with a zero charge. The discovery of these two particles gave rise to new theories regarding the intimate structure of atoms and molecules. These new theories lie within the realm of chemistry and atomic physics. Fortunately for the student of electronics, the picture of atomic structure suggested by Bohr provides a very useful physical concept for the understanding of such electronic phenomena as electron emission, ionization, and light production in gases.

A molecule is usually a combination of two or more atoms.* The normal molecule contains an equal number of electrons and protons and hence the same magnitude of negative and positive charge. If an electron is removed from a molecule, the remaining unit has an unbalanced positive charge and is called a *positive ion.* If an extra electron

* The molecules of helium, neon, argon, krypton, xenon, and radon consist of only one atom.

joins a molecule, the new unit carries a negative charge and is called a *negative ion.* The subtraction and addition of electron charges in forming ions does not change the chemical nature of the molecule since the restoration or addition of an electron will bring the molecule back to its normal neutral state. It is possible for a molecule to suffer the loss or the addition of two or more electrons. In such units the particle is called a multiple charged ion of appropriate sign. A single isolated electron is often called a *negative ion,* but this terminology will not be used in this book.

Electricity. Electron theory offers a simple explanation of the phenomenon and the properties of electricity. Such explanations may be made in terms of a displacement of electrons. Thus if one or more electrons is removed from a normal (neutral) object, a *positive charge* is created on that object. Similarly, if extra electrons are added to a neutral body, a *negative charge* is created. The magnitude of the charge is measured by the deficiency or excess of electrons from the neutral state. A common unit of charge, the coulomb, consists of 6.3×10^{18} electrons. Electric charge is represented by the symbol Q.

Electric charges may be stored in capacitors (condensers). A capacitor usually consists of two parallel conducting surfaces separated by a nonconductor. A displacement of electrons from one surface to the other causes the capacitor to be charged and energy to be stored. If the two charged surfaces are later connected by a simple conductor, the displaced electrons return to their former positions, the capacitor is discharged, and the stored energy is released.

Electric charges may exist in gases or in a near vacuum as well as on metallic surfaces. In nature, clouds are often charged negatively or positively as a result of air currents and condensation of water vapor. The charges thus acquired may be great enough to result in destructive strokes of lightning. In electron tubes either negatively or positively charged ions may become concentrated in certain regions and thus constitute a charge known as *space charge.* Space charges are important considerations in the operation of electron tubes and will be covered in later discussions.

Electrons (negative charges) and protons (positive charges) attract each other. Electrons (negative charges) repel each other. Positive ions repel each other. This basic law of attraction and repulsion of electric charges is fundamental in the operation of electronic devices.

Electric potential differences are created by a displacement of elec-

trons. If some of the electrons in a straight metal bar are moved to one end of that bar (by any means), then that end is negative and the other end is positive. A difference of potential now exists between the ends, and the magnitude of that difference depends on the density of the excess (or deficiency) of the electrons at the ends. It may also be said that electric charges exist at the ends of this rod. The magnitude of these charges depends upon the total number of electrons displaced, whereas the difference of potential depends not on the number of electrons displaced but upon the *concentration or density* of the displaced electrons. The electric charge depends on the area of the region considered, whereas the potential difference is entirely independent of the area. Potential difference is measured by the work done in carrying a unit charge from one point to another and is independent of the path followed.

Electric conduction is the process of transferring electrons in an electric circuit. *Electric current is the coordinated movement of electrons along a conductor.* The movement may be continuous, as in direct current, or periodically changing in character, as in alternating, oscillating, or pulsating currents. The magnitude of the electric current is measured by the number of electrons that move past a point in the circuit per second. An ampere of current exists when transfer takes place at the rate of 1 coulomb (charge) or 6.3×10^{18} electrons per second. Individual electrons may move at a snail's pace in a conductor of high resistance or at a speed approaching that of light in vacuum under a very high potential. The movement of individual electrons should not be confused with the propagation of an electric wave along a conductor, which takes place at a rate approaching the speed of light. The movement of electrons called *electron current* is opposite to the conventional direction of current adopted long before the electron theory was evolved. Since electronics is a science of electron movement, the direction of current flow used in this textbook for explaining tube operation will be that of the electron current.

Electric conduction takes place in solids, in liquids, in gases, and in vacuum. In solids (metals) conduction takes place through the medium of so-called free electrons. *Free electrons* may be conceived as those electrons (1) that, while forming a part of the molecule of the conductor, lie in outer orbits and thus are not very strongly bound to the nuclei, or (2) that, at instants in their orbital movement, lie midway between the nuclei of two different molecules and hence are subject to equal but opposite attractions so that they may be readily

moved by an appropriate electric field. Free electrons are moved from the sphere of one molecule subsequently to join another molecule, and thus the billions of billions of electrons in a solid conductor may move on in a kind of relay race, constituting a transfer of electrons, or electron current.

In liquids conduction is through the medium of ions. If salt (NaCl), for example, is added to distilled water, it will dissolve and disassociate into fragments or ions. The sodium atom becomes the positive ion, and the chlorine becomes the negative ion. Whether the charges exist before or only after entering solution is a question of chemistry, not electronics. If two electrodes are placed in the electrolyte and connected to a source of potential, as in Fig. 2, the ions will start to migrate to the electrode of opposite polarity. When the positive sodium ion reaches the negative electrode, it will take on an electron and will become a neutral sodium atom. Likewise, the negative chlorine ion will give up one electron on reaching the positive electrode and will become a neutral chlorine atom. The result of the migration of the two ions to the electrodes has been to transfer (in effect) one electron from the negative electrode to the positive electrode, and this transfer constitutes an electron current. The formation of the neutral sodium and chlorine atom may result in subsequent chemical actions with the water present, but these reactions are foreign to the process of electric conduction in the electrolyte.

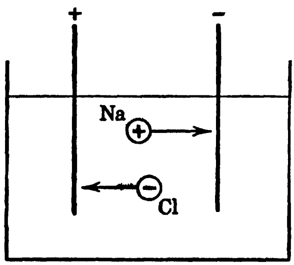


FIG. 2. Electric conduction in a solution of NaCl and water.

Electron Ballistics. The theory of the movements of charged particles in electrostatic and magnetic fields is known as electron ballistics. The movements of such particles depend upon the charges and masses of the particles, the strength of the fields, and the laws of motion.

The particle of major interest in electron tubes is the electron itself. In gaseous and vapor tubes, positive ions as well as electrons are of interest. The monatomic elements of the group of inert gases and of mercury vapor are commonly used in electron tubes. In each of these, the molecule consists of one atom. Hence, either the term atom or molecule may be used in discussing the theory of the formation and movement of ions. Table 1 gives the relative electric charges and masses of several particles existing in electron tubes.

TABLE 1 *

Name	Charge	Mass
Electron	$-e$	m_0
Positron	$+e$	m_0
Proton (H ion)	$+e$	$1{,}840m_0$
Neutron	0	$1{,}840m_0$
Alpha particle (He $^{++}$)	$+2e$	$7{,}360m_0$
Neon (ion)	$+e$	$37{,}200m_0$
Argon (ion)	$+e$	$73{,}600m_0$
Mercury (ion)	$+e$	$372{,}000m_0$

* For more precise values of m_0 and e, see R. T. Birge, "A New Table of Values of General Physical Constants," *Rev. Modern Phys.* **13** (October, 1941).

The symbol for the charge on the electron is e and for its mass m_0. Close values for the magnitude of these symbols in mks units * are as follows:

$$e = 1.6 \times 10^{-19} \text{ coulomb} \tag{1}$$

$$m_0 = 9.1 \times 10^{-31} \text{ kilogram} \tag{2}$$

$$\frac{e}{m_0} = 1.76 \times 10^{11} \text{ coulombs per kilogram} \tag{3}$$

The value given here for m_0 is for the electron moving with speeds small compared with the speed of light. For higher speeds (above 15 per cent of the velocity of light), the following expression by Lorentz should be used:

$$m_v = \frac{m_0}{\sqrt{1 - \left(\dfrac{v}{c}\right)^2}} \tag{4}$$

where m_v = mass of the electron in motion (relativity mass).
m_0 = mass of the electron at rest.
v = speed of the electron.
c = speed of light (3×10^8) meters per second.

An **electrostatic field** is a region that exerts a force upon an electric charge. The direction of the field is that in which a positive charge is urged. An electrostatic field is produced by a change of potential E with distance. The rate of change of potential and its sign is called the potential gradient. If a potential E_1 exists at one point and changes

* The mks (meter-kilogram-second) system employs the practical electric units, the volt, the coulomb, etc., making it especially useful in calculations involving electrostatic fields.

uniformly over a distance s to a value of E_2, the magnitude of the potential gradient is $(E_1 - E_2)/s$. It is customary to designate an electrostatic field by the term electric field intensity and the symbol $\mathcal{E}$. *Electric field intensity* is a vector quantity that indicates the direction of the field and its magnitude as measured by the potential gradient.

$$\mathcal{E} = -\frac{dE}{ds}^{*} \tag{5}$$

The strength of an electrostatic field is defined as the force exerted upon a unit charge. Hence, for a charge Q, the force will be:

$$f = \mathcal{E}Q \tag{6}$$

From the well-known laws of mechanics the following series of equations result:

$$f = ma \qquad \mathcal{E}Q = ma \qquad a = \frac{\mathcal{E}Q}{m} \tag{7}$$

where f is the force and a is the linear acceleration.

Since electric potential is measured by the work W done in moving a unit charge, it follows that for a charge Q

$$E = \frac{W}{Q} \quad \text{and} \quad W = EQ$$

Also,

$$\text{Potential energy (stored)} = \text{work} = EQ$$

$$\text{Kinetic energy} = \tfrac{1}{2}mv^2$$

$$\text{Kinetic energy gained} = \text{potential energy lost}$$

$$\tfrac{1}{2}mv^2 = EQ \tag{8}$$

$$v = \sqrt{2\frac{Q}{m}E} \tag{9}$$

Equation 9 shows that *both the speed and the kinetic energy acquired by a charged particle moving in an electric field is determined solely by the total potential* E *through which the particle has moved.* This fact gives

* By definition, the sign of the potential gradient is negative. In the discussion and problems that follow, the negative sign will be dropped because the interest will be in magnitude only.

the basis for a useful unit (electron-volt) for measuring the energies involved in particle motion in electric fields. The *electron-volt* is the energy acquired by an electron starting from rest and moving in a vacuum through a potential difference of 1 volt. The abbreviation for electron-volt is ev. The energy involved in electron-volts follows from equation 8.

$$\begin{aligned} &\text{1 electron falling through 1 volt} = \text{1 ev (energy)} \\ &\text{1 electron falling through } E \text{ volts} = E \text{ ev (energy)} \end{aligned} \qquad (10)$$

Also

$$1 \text{ ev} = 1.6 \times 10^{-19} \text{ joule}$$

A uniform electric field may be produced between parallel surfaces. Let A and B represent two parallel plates separated by the distance s in Fig. 3. A potential difference E applied to the plates will set up a uniform field along the line xy and also in most of the region between the plates if the dimension of the plates is large compared to the separation s.

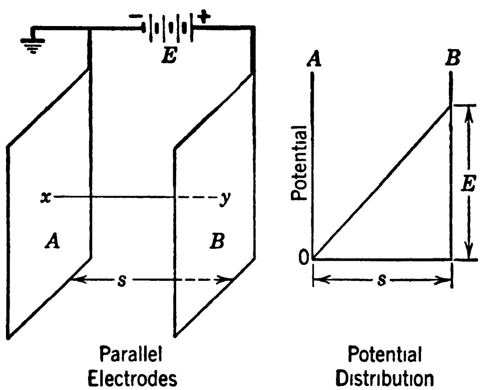


FIG. 3. Potential distribution between two planes.

The potential distribution along the line xy is shown in the right-hand view of Fig. 3. The potential rises uniformly (linearly) from 0 to E between the plates A and B so that the potential gradient and the electric-field intensity is constant in magnitude and equal to E/s. An electron (negative charge) released at x will be urged toward y with a constant force equal to $\mathcal{E}e$ and accelerated uniformly until it reaches plate B at y. If E is made 500 volts and $s = 2.5$ centimeters, the following values of motion will result (mks system).

$$\text{Potential gradient} = \mathcal{E} = \frac{E}{s} = \frac{500}{2\ 5 \times 10^{-2}} = 20{,}000 \frac{\text{volts}}{\text{meter}}$$

$$\begin{array}{l} \text{Acceleration} \\ \text{(equation 7)} \end{array} = a = \mathcal{E}\frac{Q}{m} = \frac{E}{s} \times \frac{e}{m_0} = 20{,}000 \times 1\ 76 \times 10^{11}$$

$$= 3\ 52 \times 10^{15} \text{ meters per second per second}$$

$$\begin{array}{l} \text{Speed at } y \\ \text{(equation 9)} \end{array} = \sqrt{2\frac{Q}{m}E} = \sqrt{2\frac{e}{m_0}E}$$

$$= \sqrt{2 \times 1\ 76 \times 10^{11} \times 500} = \sqrt{1\ 76 \times 10^{14}}$$

$$= 1\ 325 \times 10^{7} \frac{\text{meters}}{\text{second}} = 4\ 4 \text{ per cent velocity of light}$$

An electron beam (a pencil of flying electrons) is controlled and deflected by electric fields in cathode-ray tubes To analyze such action, assume in Fig 4 that an electron e moving with a horizontal speed of

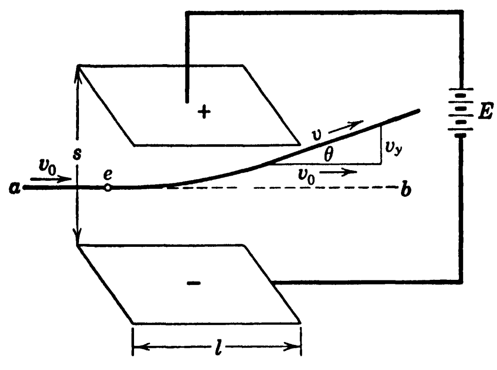


FIG 4 Deflection of an electron in an electric field

v_0 along the line aeb enters a uniform electric field (vertical) between two parallel plates at point e Assuming the polarities indicated, it is evident that the electron will be deflected upward by the electric field and travel along the full line as shown The speed of translation v_0 to the right will remain unchanged during the electron flight between the plates because the deflecting field is at right angles to that motion The electric field between the plates (assumed uniform) will give the electron an accelerated motion upward, reaching an upward speed of v_y at the time the electron has emerged from the right-hand edge of the

plates. From this point the electron will move on a straight line with a resultant speed determined by the two components v_0 and v_y. The angle θ of the resultant deflection will be determined by the component velocities v_0 and v_y.

The direction and magnitude of the resulting motion can be readily calculated as follows. Assume that the initial velocity of the electron is the same as that of the previous problem (1.325×10^7 meters per second), the plates are 1 centimeter apart with 100 volts applied, and the length of the plates l is 2.5 centimeters. The transit time for the electron having an initial horizontal component speed of v_0 is derived from

$$\underset{\text{(distance)}}{l} = \underset{\text{(speed)}}{v_0} \times \underset{\text{(time)}}{t}$$

$$t = \frac{l}{v_0} = \frac{2.5 \times 10^{-2}}{1.325 \times 10^7} = 1.885 \times 10^{-9} \text{ second}$$

$$\text{Acceleration upwards} = a = \frac{E}{s}\frac{e}{m_0}$$

$$= \frac{100}{1 \times 10^{-2}} \times 1.76 \times 10^{11}$$

$$= 1.76 \times 10^{15} \frac{\text{meters}}{\text{second}^2}$$

$$v_y = at = 1.76 \times 10^{15} \times 1.885 \times 10^{-9}$$

$$= 3.32 \times 10^6 \frac{\text{meters}}{\text{second}}$$

$$\text{Angle of deflection } \theta = \tan^{-1}\frac{v_y}{v_0} = \tan^{-1}\frac{3.32 \times 10^6}{1.325 \times 10^7}$$

$$= \tan^{-1} 0.25$$

$$= 14°$$

$$\text{Final speed } v = \sqrt{v_0^2 + v_y^2}$$

$$= \sqrt{\overline{1.325}^2\, 10^{14} + \overline{0.332}^2 \times 10^{14}}$$

$$= 1.368 \times 10^7 \frac{\text{meters}}{\text{second}}$$

The upward deflection of the electron while it is passing between the plates may be computed by using the formula

$$s = \tfrac{1}{2}at^2$$

$$s = \tfrac{1}{2}(1.76 \times 10^{15})\ (1.885 \times 10^{-9})^2$$

$$= 3.12 \times 10^{-3} \text{ meter} = 0.312 \text{ centimeter}$$

Magnetic fields are used for controlling the motion of charged particles. The action of a magnetic field upon charged particles follows from the theory of the force acting upon a current-bearing conductor placed in a magnetic field.

$$f = \beta i \sin \theta$$

where f represents force per unit length, β the flux density, i the current, and θ the angle the conductor makes with the magnetic field. A current consists of electrons or other charged particles in motion. Hence

$$i = Qnv$$

where Q is the charge on the particle, v is the speed of the particle, and n is the number of particles per unit length. Since the force on the individual particle is under consideration, n becomes unity and

$$f = \beta Qv \sin \theta$$

In Fig. 5, let it be assumed that an electron e is projected at right angles ($\theta = 90°$) into a uniform magnetic field with an initial velocity of v_0. Since the field is directed in (toward the paper), the application of any convenient rule * for the force upon a conductor will show that the electron will be deflected downward as it enters the field. With a uniform field and a constant speed v_0, the path of the electron will be given a constant angular change or acceleration and will move in the arc of a circle. From the laws of mechanics this acceleration is v^2/r, where r is the radius of curvature of the path. Thus

$$f = ma = \beta Qv = m\frac{v^2}{r}$$

and

$$r = \frac{mv}{\beta Q} \tag{11}$$

and for an electron where $Q = e$

* Use left hand instead of right hand for electron current.

$$r = \frac{v}{\beta} \times \frac{1}{e/m} = 5.69 \times 10^{-12} \frac{v}{\beta} \text{ meters} \tag{12}$$

Assume in Fig. 5 that v_0 is 2×10^9 centimeters per second and the field strength β is 10 gauss.

$$r = 5.69 \times 10^{-12} \frac{\overset{\text{(meters per second)}}{2 \times 10^7}}{\underset{\text{(webers* per square meter)}}{10 \times 10^4 \times 10^{-8}}}$$

$$= 11.38 \times 10^{-2} \text{ meter} = 11.38 \text{ centimeters}$$

The theory of the magnetic field at right angles to the moving electron (illustrated in Fig. 5) is applied in the deflection of a beam of electrons and is analogous to the use of the electric field as illustrated in Fig. 4.

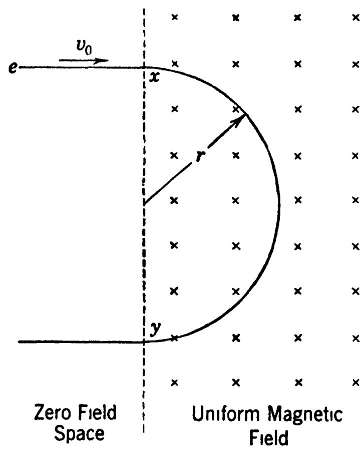


FIG. 5 Circular deflection of an electron in a uniform magnetic field

The angle of deflection of an electron beam in a uniform magnetic field may be computed from the relations shown in Fig. 6. Here a magnetic field within the rectangle *abcd* is created by two parallel coils, one in front of and one in back of the page. An electron entering the field from the left with a velocity v_0 at point x will be deflected along the arc of a circle xy having some radius r. After emerging from the field at y, the electron will continue its velocity v_0 along a linear path. The angle of the arc xy is θ (lower angle) and by plane geometry is equal to the angle of deflection θ (at the right). Obviously, the sine of the angle θ is l/r. The length of field l is known and r can be calculated from equation 12 as in the preceding example.

If a flying electron enters a magnetic field at an angle less than a right angle, an interesting phenomenon occurs. This is illustrated in Fig. 7. An electron moving with a velocity v_0 enters a horizontal magnetic field at point e in a direction making an angle θ with the x axis. The horizontal magnetic field is produced by the solenoid $CCC'C'$.

* One weber equals 10^8 maxwells.

Assume for the sake of a simple concept that the magnetic field begins at the boundary CC and ends at boundary $C'C'$. The initial velocity of the electron v_0 has a horizontal component v_x and a vertical component v_y. The horizontal component v_x is parallel to and unaffected by the magnetic field. The vertical component v_y is normal to the magnetic field and hence is subject to a side thrust which will result in a circular motion as explained in the preceding discussion. The resultant motion of the electron while it is in the magnetic field will be a combination of a motion of translation to the right equal to v_x and a circular motion about a horizontal axis. Such a combination of motions follow a helical path as shown in Fig. 7.

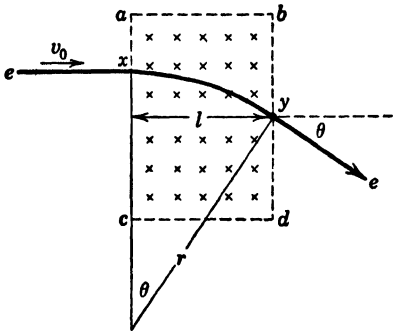


FIG. 6. Angle of deflection of an electron traversing a magnetic field.

It is obvious that the effect of the magnetic field has been to change the direction of motion of v_0 into a helical motion along the direction of the field or x axis. Accordingly, a beam of electrons projected into a magnetic field at various angles other than 90 degrees will be deflected into the direction of the magnetic field. This is the principle

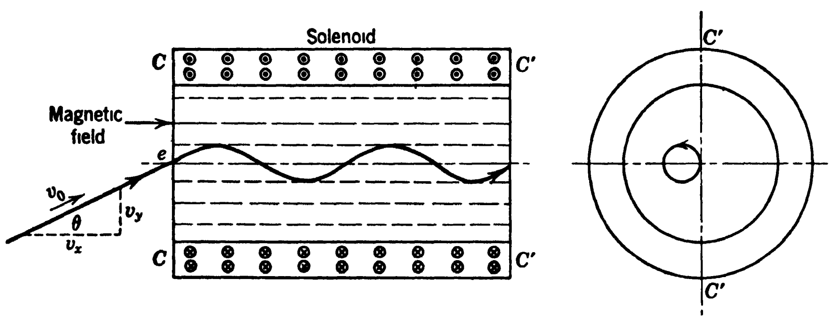


FIG. 7. Helical path of an electron after entering a longitudinal magnetic field in a solenoid (magnetic focusing).

of magnetic focusing which is applied in some cathode-ray tubes and electron microscopes. In the application of one form of magnetic focusing, the field strength is adjusted so that the electrons make one or an integral number of turns in the helical path. If all electrons

enter on the axis of the magnetic field (*though at various small angles*) and make the same number of turns, they will converge on one spot. Calculations for the path of the electron in the magnetic field can be

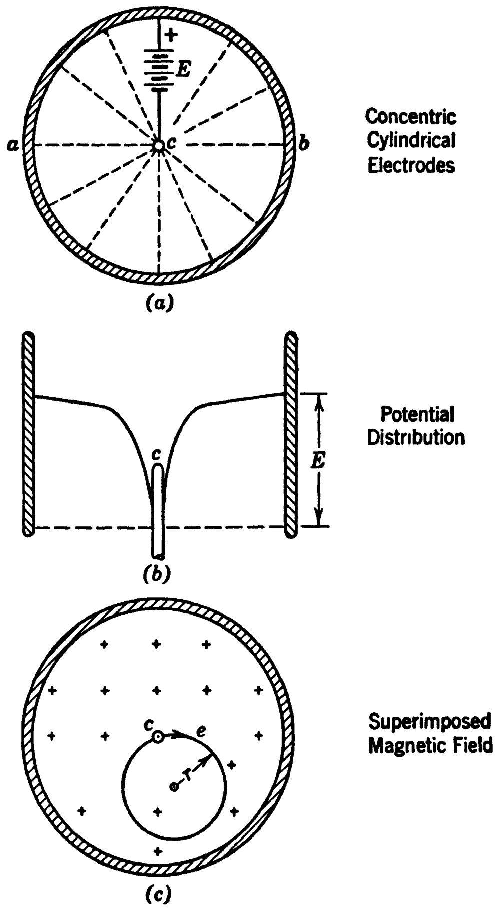


FIG. 8. Approximate path of an electron in a combined electric and magnetic field between concentric cylindrical electrodes.

made by adapting the theory of the circular motion to the v_y component of v_0 and using the physical concept of the helix.

Concentric cylindrical electrodes are frequently used in electron tubes. A common configuration is shown in part *a* for Fig. 8. A potential E is applied to the cylinders so that the central electrode c is negative and the resulting electric field will be directed along radial

lines as shown (dotted lines). In practice, electrode c may be a very fine wire giving a potential distribution along any diameter as shown in part b of Fig. 8. The change in potential or potential gradient close to c is very great and then becomes slight throughout most of the radial path to the outer cylinder. If electrons are released at the wire c, they will fly along radial lines under the influence of the electric field to the outer cylinder.

Combined electric and magnetic fields applied to the configuration discussed in the preceding paragraph result in an interesting and useful phenomenon. If a solenoid bearing continuous current is placed around the concentric cylindrical electrodes, a field parallel to the axis of the cylinders will be created as indicated by the crosses on part c of Fig. 8. Now the electrons released at c move outward under the influence of the electric field as before, but in doing so their path is normal to the magnetic field so that they are subject to a deflection in a curved path. If the magnetic field is sufficiently strong, these electrons may never reach the outer electrode but may be returned to electrode c as shown in the figure. This phenomenon has been utilized in a tube known as a magnetron and applied in an oscillator for producing very high frequencies.

In one application of the magnetron, the magnetic field is adjusted so that the electron will just graze the outer cylindrical electrode in following its curved path. The exact mathematical solution for this motion is too involved to justify treatment here.

Electron optics is the science of controlling electron beams in a manner analogous to the control of light rays by optical lenses. The electron beams are directed in parallel, converging, or diverging paths through the medium of electric and magnetic fields. The focusing effect of a parallel magnetic field illustrated in Fig. 7 is an example of the methods employed. A converging or a diverging magnetic field will exercise a corresponding directive effect upon an electron beam. An electric field between adjacent concentric cylinders may be used to produce a converging or diverging effect upon a movement of electrons. This phenomenon is illustrated in Fig. 9 where the electric-field distribution shown is created in the proximity of the adjoining cylinders. Electrons moving along the general direction of the axis of the cylinders will tend to move parallel to the lines of the electric field. An electron such as e moving along the line AB in Fig. 9 will be deflected as shown, and a group or beam of diverging electrons will be directed so as to converge at a point X on the axis. By such methods and com-

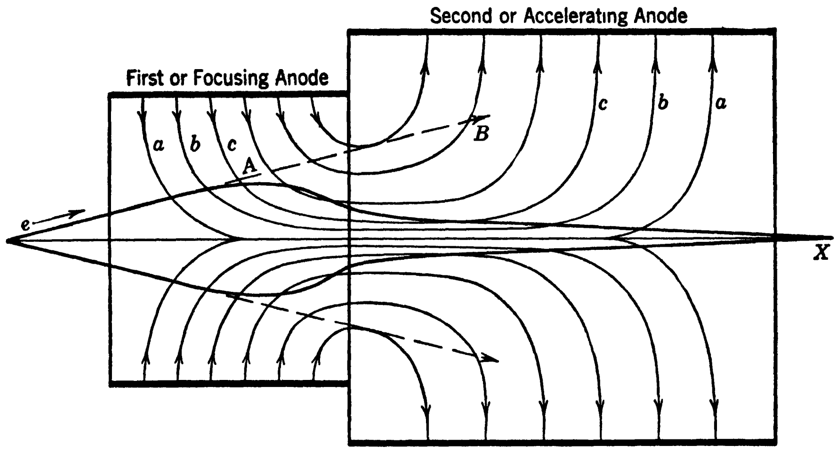


FIG. 9. Focusing of an electron beam by an electrostatic lens.

binations of electric and magnetic fields, the art of electron optics is accomplished. The complete theory of electron optics is a subject too advanced for the purpose of this book.

PROBLEMS

1. Two large parallel planes in a high vacuum separated by a distance of 1 inch carry a difference of potential of 1000 volts. An electron is freed at the negative plane (near center) with zero velocity. Calculate (a) the potential gradient between the planes, (b) the velocity of the electron when it hits the positive plane, (c) the time the electron is in transit, and (d) the energy acquired by the electron in electron volts and joules (use mks system).

2. Increase the distance between the planes of Problem 1 to 5.08 centimeters, and recalculate.

3. (a) In Problem 1 substitute a + hydrogen ion (at positive plate) for the electron and solve. Repeat for (b) a neon ion, and (c) a mercury ion.

4. In Fig. 4, an electron enters the space between the two planes with an initial velocity of 10^9 centimeters per second. If the electric field between the planes is 5 volts per millimeter, what will be the angle of flight and the resultant velocity after the electron has traveled 2 centimeters to the right?

5. In Fig. 5, an electron is hurled with a velocity of 10^9 centimeters per second into a magnetic field of 15 gauss. Calculate the nature of the path of the electron. What will be the resultant velocity 1/1,000,000 of a second after the electron enters the magnetic field?

6. If a second magnetic field of equal strength is added at right angles (directed top to bottom) to the field of Problem 5, what will be the nature

of the electron path? Will the resultant velocity change under these conditions?

7. An electron falling from rest in a uniform electric field acquires a velocity of 3.5×10^9 centimeters per second. What is its energy in joules? Through what potential has it fallen?

8. Substitute an argon ion with a velocity of 10^6 centimeters per second for the electron in Problem 7 and solve.

Chapter II

ELECTRON EMISSION

Electron emission is the liberation of electrons from the atomic forces in metals. Many of the physical phenomena connected with the escape of electrons were observed in the nineteenth century. In 1827 Robert Brown observed minute particles of dead matter in suspension in water under a high-powered microscope and discovered that they performed irregular wiggling motions, suggesting "life." This phenomenon came to be known as the *Brownian movement* and was explained a half century later. The Brownian movement is due to the continual bombardment of inanimate particles by the thermal agitation of the molecules in the liquid. Similar movements take place in gases and in solids. In 1883 Thomas Edison observed that the region surrounding a red-hot filament is a conductor of electricity. He placed a metallic plate in an incandescent (carbon) lamp bulb and connected it in series with a galvanometer. His observations showed that whenever the plate was connected to the positive side of the filament a current was indicated, but when it was connected to the negative side there was no deflection of the galvanometer. This phenomenon came to be known as the *Edison effect*. In 1888 Hallwachs discovered that, if he charged a zinc plate to a negative potential and then exposed it to ultraviolet light, it gradually lost its charge. In each of the experiments cited, electrons were freed from the bonds of electron affinity.

Electron Affinity. Electrons are bound within the surface of a metal by an attraction or affinity. A concept of this affinity may be attained by reference to the mechanical picture of the Bohr atom. First, one may visualize an isolated atom with its positive nucleus and the electrons revolving around it in a series of orbits at various energy levels. These electrons are bound to the nucleus by an attractive force which gives them potential energy. In their orbital motion they have kinetic energy which balances the potential energy. The electrons in the outer orbit (the valence electrons) are less closely bound and are the ones that become free electrons and participate in electric conduc-

tion and in electron liberation. Individual atoms combine into molecules held together by a displacement of the valence electrons. The molecules, in turn, are held together in the solid by a rearrangement of electrons in a form of lattice structure. Thus the electrons that may participate in any liberation are subject to an attraction or affinity due to atomic structure, molecular structure, and molecular combinations. The Bohr atom may be envisioned within a section of metal as shown in Fig. 1. An electron existing within the metal at some point such as Y will be subject to many attractions from its parent atom and from surrounding atoms. This electron will have an absolute motion consisting of its private orbital speed plus the motion of its parent atom due to thermal agitation. Obviously, the electron will occupy many positions in the course of a fraction of a second where the forces of attraction acting upon it will be balanced. While in this balanced state it is "free" in a sense, and any electric field that is present will cause it to move and to become a part of electric conduction. However, an electron in position X at the surface of the solid never reaches any balanced state but, on the contrary, is subject to all the atomic and molecular forces that bind it to the solid. This combination of forces is electron affinity. Electron affinity produces a strong potential barrier which makes it very difficult for an electron to break away from the surface of its parent solid. In order to break away, an electron must have an initial velocity or kinetic energy sufficient to overcome the barrier or electron affinity. This velocity is the resultant of absolute motion of the electron arising from the combination of electron (orbital) and thermal atomic movements. If an electron carrying a negative charge e succeeds in escaping from the surface to some position such as Z, it leaves behind a positive charge e which exerts a powerful force to bring it back to the metal. This positive charge which remains in the metal is called the *image positive charge.*

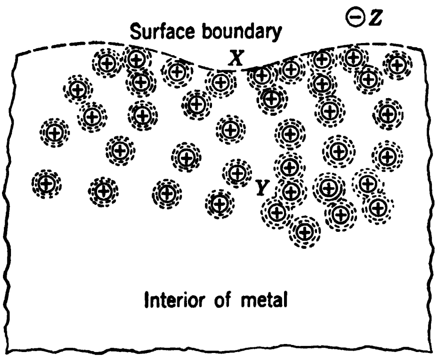


FIG. 1

The preceding discussion indicates that energy or work is required to remove electrons from solids. This energy for removing a single

electron from a solid may be designated as w. The ratio of this work w to the charge on the electron e is called the *work function* of the material. This ratio w/e is the same as the ratio of work to unit charge which, in turn, is the definition for the potential difference between two points. Thus it appears that work function or *work function equivalent* can be expressed in volts. A second way to view this subject of work function is to remember that an electron escapes from a solid through its kinetic energy. Thus

$$w = \tfrac{1}{2}mv^2$$

and the velocity v could be acquired by the action of a potential E (volts) acting on the electron with charge e. Hence

$$w = Ee = \tfrac{1}{2}mv^2$$

and work function $= w/e = E$.

To develop an understanding of work function, the student should keep in mind the following three statements. (1) Work function is a measure of the work required to overcome electron affinity plus that required to overcome the positive image charge on the surface after the electron escapes. (2) The energy of the work function is expressed in volts (electron volts), but this *does not mean* that a positive potential equal to the work function placed outside the surface will extract electrons. (3) The values of work function given in this text and in the literature refer to that *additional energy* (above that possessed at normal temperature) which is necessary to free them from the parent surface. The work function in volts for a number of materials used in electron tubes is given in Table 1. Additional discussion concerning work function will be given at the end of the chapter to explain some characteristics of the materials commonly used in thermionic emission.

TABLE 1

Material	Work Function ϕ (volts)	Material	Work Function ϕ (volts)
Platinum	5.0	Magnesium	2.7
Tungsten	4.52	Thorium on tungsten	2.63
Carbon	4.5	Calcium	2.5
Mercury	4.4	Barium	2.0
Molybdenum	4.3	Sodium	1.9
Tantalum	4.1	Calcium oxide	1.9
Nickel	4.0	Strontium oxide	1.4
Copper	4.0	Barium oxide	1.1
Thorium	3.0		

Electron Emission. Electrons may be liberated from metals in five different ways, as follows: (1) high-field emission, (2) secondary emission, (3) thermionic emission, (4) photoelectric emission, and (5) radioactive disintegration. Photoelectric emission consists of the removal of electrons by electromagnetic wave radiations such as light and will be discussed in Chapter VII. Radioactive materials eject electrons (beta rays) during their slow process of disintegration. This phenomenon is important in many scientific studies. In some electron tubes which contain gas and which start with cold cathodes, it is probable that radioactive materials within or near the tube and sometimes photoelectric emission provide the initial electron emission for exciting or starting the functioning of the device. The other methods for producing electron emission will be discussed in the articles that follow.

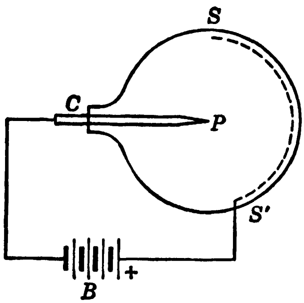


FIG. 2. An electron microscope employing high-field emission.

High-field emission is the liberation of electrons from cold metals by virtue of a very high potential gradient at the surface of the metal. At normal temperatures relatively few electrons in a metal attain a velocity at the surface sufficient to overcome the surface potential energy barrier. However, some of them do so, but upon emerging from the surface they leave behind a positive image charge which attracts them back into the surface. Accordingly, the electrons that do emerge move only an infinitesimal distance from the surface before returning. A very powerful electric field having a magnitude of over one million volts per centimeter is required to pull such electrons from the surface after this transient release. Such fields are difficult to attain, and this method finds little application in electronic devices. The student should remember the extreme difficulty of extracting electrons from cold electrodes, since this fact forms the basis of unilateral conductivity in electron tubes.

One excellent example of high-field emission has been demonstrated by J. A. Becker of the Bell Telephone Laboratories in the form of an electron microscope illustrated in Fig. 2. A spherical glass tube exhausted to a high vacuum contains a metal electrode C and a conducting fluorescent screen SS'. The electrode C is swaged to give a very fine point and is further treated by repeated dippings in an acid solution to remove any impurities and to attain a point P of nearly infinitesimal

size. A potential of 230 volts applied at B was capable of producing a potential gradient of over one million volts per centimeter at point P because of the sharpness of the point. The electrons attracted to the fluorescent screen gave a beautiful picture of the structure of the metal point P.

Another example of the use of high-field emission occurs in mercury-pool tubes where a positive space charge created by positive mercury ions produces a high field at the surface of the metallic mercury. This application will receive subsequent treatment.

Secondary emission is the ejection or "splashing-out" of electrons from a solid due to bombardment by electrons, positive ions, or other flying particles. Usually this emission is due to electrons attracted to an electrode having a positive potential. A single primary impinging electron may eject from one to ten secondary electrons from the electrode, depending upon the work function of the material, the condition of its surface, and the velocity of the primary electron. Obviously, the kinetic energy of the primary electron is imparted to the secondary electrons and added to their normal energy in such a way that they are able to overcome the potential barrier or work function of the surface. In general, it would be expected that the energy of the impinging electron must be equal to or greater than the work function of the electrode surface where pure metals are concerned. The presence of adsorbed gas in the surface of the electrode increases the secondary emission. Here the presence of the gas molecules under the surface molecules weakens the potential energy barrier and permits the kinetic energy of the impinging electron to become more effective.

Secondary emission occurs at some electrodes in nearly all electron tubes. Sometimes its presence goes unnoted, sometimes it is a disturbing factor, and in other tubes it is employed for a useful purpose. Each case will be treated in appropriate future articles.

Thermionic emission is the liberation of electrons from a metal produced by the thermal agitation of its atoms. The phenomenon is analogous to the evaporation of liquids. At normal temperature the molecules of a liquid have some thermal agitation, but few of those at the surface "jump out" far enough to remain as vapor. With a rise in temperature the individual motions of the molecules become more violent and an increasing number do overcome the attraction of the liquid and do evaporate or "boil out." In like manner, the electrons in a metal are closely held by electron affinity, and a relatively small number have sufficient thermal energy to break away from the surface at ordinary temperatures. With a rising temperature the thermal movement of

the atoms and the kinetic energy of the electrons increase so that more and more succeed in breaking through the potential energy barrier at the surface of the metal. As the electrons break out in space, the force of attraction of the positive image charge remaining on the metal soon overcomes the initial velocity of emission, and the electrons drop back into the metal. Thus, when operating in a zero electric field, the electrons never get very far from the surface of the heated metal.

The phenomenon of thermionic emission may be conceived as illustrated in Fig. 3. Let it be assumed that a filament or cylindrical conductor of tungsten is placed in a vacuum or in a space filled with inert gas at low pressure. At a certain temperature of the filament the thermal motion of the molecules and electrons becomes great enough so that a number of electrons are thrown out as shown in part a of Fig. 3. A further rise in temperature of the filament will be accompanied by an increase in the kinetic energy of the electrons and an increase in the number and initial velocity of those emitted (part b of Fig. 3). A still higher temperature of the filament will result in a greater agitation and velocity of electrons so that a cloud or atmosphere of electrons will be formed as depicted in part c of Fig. 3. It should be understood that Fig 3 represents an isolated filament and the surrounding space magnified many times Many electrons emitted from the filament never travel over 0.01 millimeter before returning, and few go farther than 1.5 millimeters. It is obvious that the initial velocity of emission will vary over wide limits, a large number having a relatively low velocity and a few having a velocity sufficient to carry them into the outer region of the electron cloud.

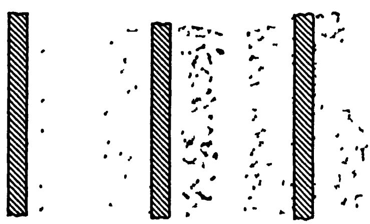


FIG. 3. Changes of thermionic emission with temperature.

The initial velocity of emission that an electron must have in order to break away from a tungsten filament can be readily calculated from equation 9, page 8, where E represents the work function of tungsten in volts (Table 1).

$$v = \sqrt{2\frac{e}{m}E} = \sqrt{2 \times 1.76 \times 10^{11} \times 4.52}$$

$$= 1.26 \times 10^6 \text{ meters per second}$$

$$= 2{,}820{,}000 \text{ miles per hour}$$

The purity of an emitting material and the cleanliness of its surface have a considerable bearing upon the amount of thermionic emission. Wilson found that the emission of a hot platinum filament may be reduced to 1/250,000 of its normal value by first heating it in oxygen or boiling it in nitric acid. A small amount of hydrogen around a heated filament overcomes the effects of oxygen and nitric acid. The presence of a small amount of water vapor in a tube will greatly reduce the thermionic emission. Special care is used in the preparation of emitters to eliminate or neutralize the harmful effects of occluded gases and water vapors. Thermionic emission may be greatly increased by a surface layer (usually one atom thick) of thorium, barium, strontium, or calcium on a base of tungsten, nickel, and some other metals.

Equation of Thermionic Emission. Early in the twentieth century O. W. Richardson reasoned that electron emission from hot solids bore a similarity to the evaporation from liquids. Using the classic kinetic theory, he developed an equation representing electron emission and performed experiments that checked the correctness of his theory to his own satisfaction. Some contemporary scientists did not obtain equally satisfactory experimental results, which led Langmuir to refine the experiment by eliminating errors resulting from the presence of gas in the tube and impurities in the filament and thus to verify the form of Richardson's equation. Richardson suggested a second equation for emission which was formulated by S. Dushman and checked experimentally. The latter equation, which holds general acceptance today, is as follows:

$$I = AT^2\epsilon^{-b_0/T} \tag{1}$$

where I = the emission current in amperes per square centimeter of emitting surface (saturation current density).

T = the absolute temperature (degrees C + 273) = Kelvin.

ϵ = the natural base of logarithms.

A = a constant.

$$b_0 = \frac{\phi}{k} = \frac{\text{Work function}}{\text{Boltzmann's constant } (863 \times 10^{-4} \text{ volt per degree K})}.$$

The various factors that control electron emission from hot bodies have made it difficult to correlate the values of constants derived from scientific deductions with those determined experimentally. Accordingly, the constants are determined by experiment and calculations, and the equation is treated as empirical.

For pure metals A is a universal constant, but for coated metal cathodes both A and b_0 vary because of the degree of coverage of coated material, temperature, and other factors. Experimental values of the constants for some cathode materials are given in Table 2.

TABLE 2

	A	b_0
Tungsten	60.2	52,400
Thoriated tungsten	3.0	30,500
Barium oxide	0.01	12,000

The curve representing the trend of Richardson's equation is given in Fig. 4. It should be noted that this curve and equation 1 give the emission current (saturation current density) from a hot body or cathode and do *not* represent the current that may pass to a plate or anode in an electron tube.

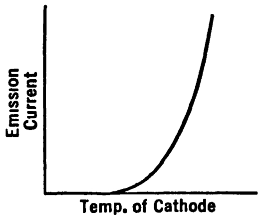


FIG. 4. Curve of thermionic emission.

Construction of Cathodes. The electrode from which electrons pass into the vacuum or gas in an electron tube is called the *cathode*. Cathodes may be classified as hot or cold. Hot cathodes may be either directly heated or indirectly heated. The directly heated type uses a filament construction as illustrated in parts *a*, *b*, and *c* of Fig. 5. The filament of parts *a* and *b* consist of a flat ribbon or round wire which is heated by current passing throughout. Part *c* shows a wire filament wound into a fine helix. The helical construc-

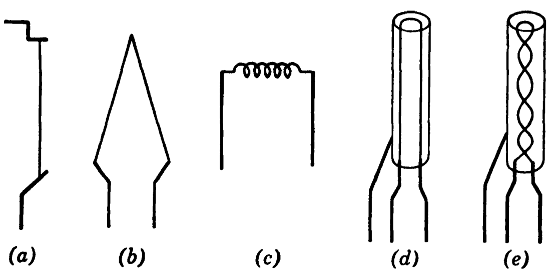


FIG. 5. Construction of typical cathodes for vacuum tubes.

tion is also used for other configurations such as in parts *a* and *b*. The material for the emitter of the filamentary cathodes of *a*, *b*, and *c*, Fig. 5, may be pure tungsten, thoriated tungsten, or an oxide-coated alloy of nickel. The cathode of the indirectly heated type is illus-

trated in parts *d* and *e* of Fig. 5. It consists of an oxide-coated sleeve or tube which constitutes the cathode proper plus a loop or twisted tungsten heating unit placed inside the sleeve. The tungsten heater is insulated from the sleeve by a coating of aluminum oxide. The cathode heater tube is usually made of nickel or some nickel alloy such as Konal—an alloy of nickel, cobalt, iron, and titanium.

Tungsten Cathode. The desirable properties of a material for a thermionic emitter of electrons are a high melting point, a low work function, and a long life. Tungsten has been used as an emitter for many years. It melts at 3600 degrees K and is generally operated at temperatures from 2450 degrees K to 2600 degrees K. At these temperatures it can be operated for several thousand hours to furnish an excellent supply of electrons. Tungsten has a relatively large work function so that its efficiency as measured in amperes of emission per watt of heating power is rather low. Pure tungsten was used in early electron tubes of all kinds and is still used in tubes that have a high plate voltage (above 10,000) and wherever severe positive ion bombardment of the cathode is likely to occur. The characteristics of pure tungsten as an emitter are shown graphically in Figs. 6 and 7.

Other metals that have been used sometimes for cathodes are tantulum, columbium, platinum, and molybdenum. Tungsten is preferred because of its mechanical strength and because it can be obtained more readily.

Thoriated-Tungsten Cathode. The thoriated-tungsten cathode was developed by Langmuir and his co-workers. This cathode is made of pure tungsten impregnated with approximately one per cent of thorium oxide (thoria). At its normal operating temperature of 2000 degrees K, the uncarbonized thoriated-tungsten cathode has an emission per unit of surface more than 10,000 times that of a pure tungsten filament. And its emission (at 2000 degrees K) is more than 90 times that of tungsten at the temperature of 2400 degrees K.

The thoriated-tungsten cathode is prepared for service by being mounted in a tube and heated to 2700 or 2800 degrees K for a few minutes. It is then operated at the normal operating temperature of 1900 to 2000 degrees K. The first heating at the higher temperature reduces part of the thorium oxide to pure thorium. The thorium atoms thus formed diffuse through the body of the cathode and slowly come to the surface where they form a skin or layer one molecule thick. Thus the emission of electrons comes from the thin layer of thorium atoms which have a low work function and hence give a copious emission. At the normal operating temperature, the thorium atoms are

evaporated slowly from the cathode but others from the inside diffuse to the surface to take their place. If the thoriated cathode is raised to a higher temperature, the electron emission increases but the rate of evaporation of thorium also increases and, at a certain critical temperature of about 2200 degrees K, all the thorium atoms leave the surface. Since the diffusion from the inside cannot be rapid enough to replace

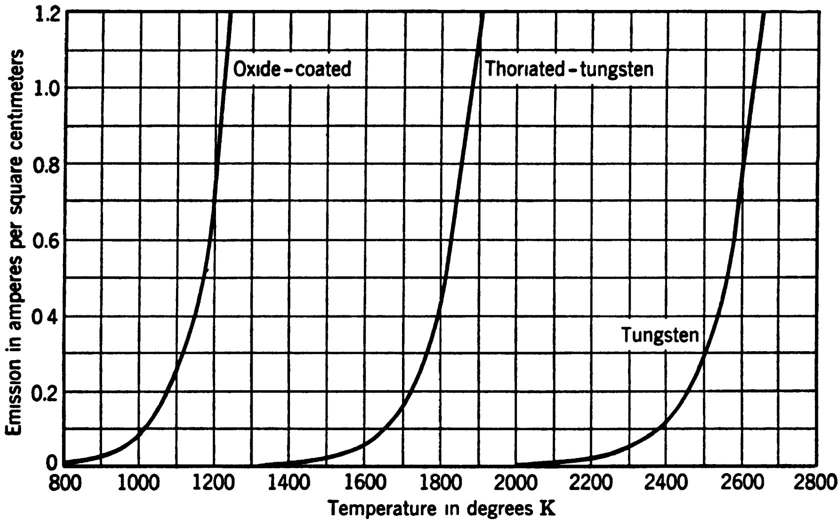


FIG. 6. Emission of typical hot-cathode emitters.

them, the emission drops to the value for pure tungsten. Thus the temperature of operation of this emitter is important in order to maintain the proper equilibrium between the thorium supplied from the interior and the amount distilled from the surface.

After thousands of hours of service, all the thorium atoms within the cathode may be used for replacements and then the emission of the cathode will fall off. In general, the filament or cathode may be rejuvenated by heating it for a short time to 2800 degrees K and then by restoring the temperature to 2000 degrees K. The first step reduces more thorium oxide to thorium and probably drives from the filament certain impurities, such as gas atoms, which have become occluded in it during the period of operation. This step is sometimes called deactivation since it stops the emission of the thorium for the time being. The second step is known as activation. Under activation the normal diffusion of thorium restores the thorium atomic layer on the surface and within a few minutes the emission returns to its full value. The

quantity of thorium oxide in the cathode is usually sufficient to repeat the deactivation and activation process several times.

The thoriated-tungsten cathode is sensitive to bombardment by positive gas ions which knock off the thorium surface atoms and reduce the

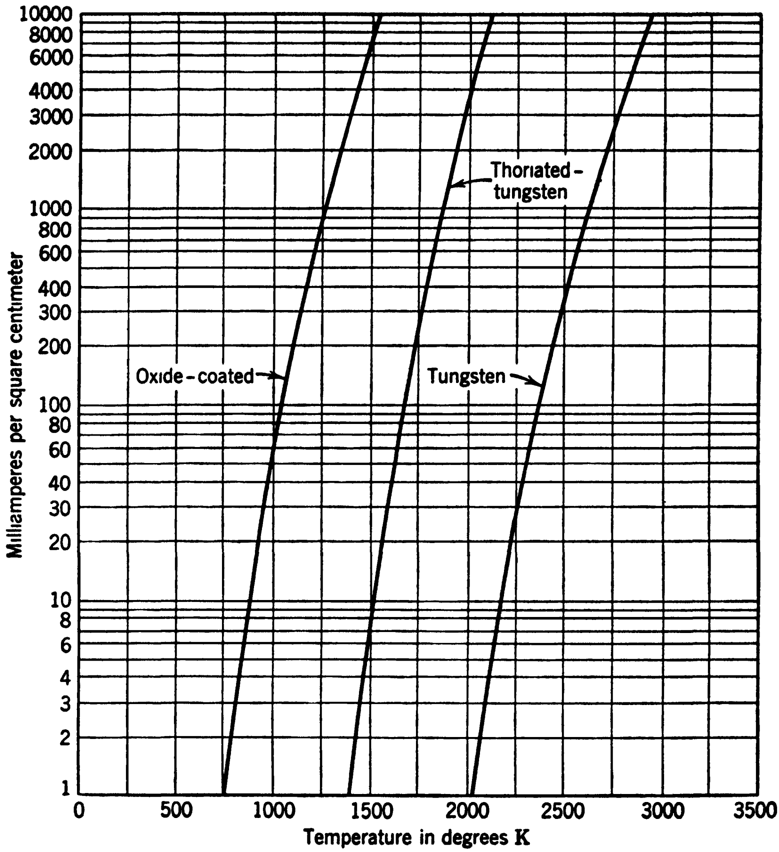


FIG. 7. Thermionic emission of cathodes.

emission. This sensitivity has been reduced by a carbonizing treatment which is given to all thoriated-tungsten cathodes manufactured today. Two methods have been employed for carbonizing. In one method the cathode is placed in a vacuum for cleaning; acetylene gas under low pressure is then introduced and the cathode is "flashed" at a temperature of 2200 to 2300 degrees K. In this process the surface of the cathode is reduced to tungsten carbide, W_2C, which has a higher resistance than pure tungsten. During "flashing," the cathode is held

at a constant impressed voltage until the current decreases by 7 to 10 per cent of its original value. This decrease in current occurs when 20 to 30 per cent of the cross section of the cathode has been converted to W_2C. Further carbonization is undesirable because W_2C is brittle and weakens the cathode to mechanical shock. Carbonization at a temperature lower than that suggested results in the formation of WC, which is unsatisfactory for the desired purpose. Carbonization can be effected by cleaning the cathodes in hydrogen vapor and then flashing at atmospheric pressure in nitrogen containing a low percentage of hydrocarbon vapor such as hexane or benzene, with the same control as suggested previously. While 4000 volts has often been suggested as an upper limit for plate voltage for thoriated tungsten, successful operation up to 10,000 and even 15,000 volts has been secured in carefully evacuated transmitting tubes. Normally, the carbonized cathode does not require any additional treatment for activation; when treatment is needed, it is the same as suggested earlier.

The emission characteristic of the thoriated-tungsten cathode is illustrated in the curves of Figs. 6 and 7. This emitting material (uncarbonized) was widely used in the construction of tubes for radio receiving sets during the period 1925–1930. The general adoption of the heater type of cathode after that period has confined the application of the thoriated-tungsten cathode to power amplifier tubes operating in the range of 750 to 5000 volts.

Oxide-Coated Cathode. In 1904 Wehnelt discovered that a small amount of calcium oxide on platinum greatly increased the electron emission. This discovery led to many experiments using various alkaline earth oxides placed on metallic cores and to the development of excellent cathodes. Barium and strontium oxide were found to be excellent materials; a mixture of approximately 50 per cent barium oxide and 50 per cent strontium oxide provides a coating with satisfactory mechanical properties and a copious emission.

The oxide-coated cathode is analogous to the thoriated-tungsten cathode in a number of ways. First, the addition of the oxide increases the emission as does the thorium coating. Second, the oxide-coated cathode has a high emitting efficiency like the thoriated-tungsten cathode. Third, any overheating of the oxide-coated filament above a critical temperature greatly reduces the emission. Fourth, the oxide-coated filament must be activated before it is ready for service.

The base or core for the oxide-coated cathode is generally made of nickel because of its good physical properties and low cost. Several alloys have been developed for the cathode base, among which Konal

offers a speedy chemical reduction of the oxide, giving a supply of the free alkaline earth metal for activation. The alkaline earth oxides are not stable in ordinary air; hence it is necessary to coat the core with their carbonates, nitrates, or hydroxides and subsequently to reduce these to oxides. The barium and strontium carbonates or nitrates may be applied to the emitter base by dipping the filament in a water suspension of the salts or by spraying the base with a suspension in any acetate with a little nitrocellulose for a binder. For tungsten filaments, hydroxide coatings may be applied by dipping the filaments in a bath of molten hydroxide. After the coating is applied, emitters are dried in air by heating.

The prepared filament is now mounted in the tube in which it is to be used. The tube is evacuated and kept on the pumps while the cathode is heated to a temperature of about 1400 degrees K to reduce the coating to oxide. Unless some powerful reducing agent has been present, the filament will still be "unactivated." This means that its electron emission at the working temperature of 800 to 1000 degrees K is still very low. Activation is now brought about by (1) prolonged heating, (2) applying a potential of several hundred volts to the anode, or (3) both. When the electron emission becomes about 300 milliamperes per square centimeter for a temperature of 1100 degrees K, there will be no further increase and the filament is ready for use. Activation means that free barium or strontium has been liberated from the oxide by chemical action, by electrolysis, by positive ion bombardment, or by a combination of these processes. The free metal atoms are distributed throughout the body of the oxide and on its surface in such a way as to induce a copious electron emission from the outer surface. It is generally agreed that the emission comes from the metallic atoms and not from the oxide or the base of the cathode.

After several thousand hours of normal use the emission from the oxide-coated cathode declines rather rapidly. When this point is reached the tube should be discarded since little can be done to restore its emitting property. Tubes using oxide-coated cathodes are frequently guaranteed for a life of 8000 hours and commonly operate for 20,000 hours and longer. The temperature emission characteristics of oxide-coated cathodes are shown in Figs. 6 and 7.

Oxide-coated cathodes are made in both the filament and heater types illustrated in Fig. 5. The operating temperatures for these cathodes vary from about 800 to 1100 degrees, depending upon the emission required, the expected life, and the application.

Effect of Gas upon Emission. The presence of gases has a harmful effect upon the emission and operation of *vacuum* tubes. It has been pointed out that oxygen and water vapor greatly reduce the emission from some materials and that the bombardment by positive ions may destroy the emitting surface layer of coated cathodes. Thus it is imperative that harmful gases be removed and, in the case of vacuum tubes, that all gases be removed as far as feasible. The gas that is present in a tube comes from two sources. First, there is the gas that occupies the open space in the tube, and, second, the gas that is occluded or adsorbed in the surface of the metal, glass, and other materials inside the tube. The occluded or adsorbed gas can be removed by heating the parts where it resides. This heating may be accomplished by passing current through the material as in emitters, by placing the tube in a high-frequency induction field, or by electron bombardment of electrodes. Oxygen may be removed by using hydrogen gas or by exhausting the tube and then refilling the space with an inert gas. The final evacuation is accomplished in two ways. First, a vacuum pump removes a large part of the gas, and then a material known as a "getter" completes the operation. The getter is a chemical substance, such as barium, magnesium, aluminum, and tantulum, that has the property of combining with gases when they are vaporized. In glass tubes a small amount of getter is mounted in a position where it will be heated and vaporized by the induction field. The effectiveness of the getter results from its chemical combination with the gases during flashing and from subsequent absorption of the gases by the residue of the getter which is deposited on the walls of the tube. The methods for the removal of gas are usually combined in various ways in the manufacture of vacuum tubes.

Comparison of Emitter Materials. The characteristics of emitters vary so widely with the temperature and other factors that it is difficult to make comparisons. An approximate comparison of emitting

TABLE 3

Cathode Material	Temperature (Degrees K)	Amperes per Square Centimeter	Milli-Amperes per Watt	Watts Required per Ampere
Tungsten	2450–2600	0.2–0.65	3–8	333–125
Thoriated-tungsten	1900	1.15	100	10
Oxide-coated	800–1100	0.01–0.22	50–300	20–3.3

efficiency is given in Table 3. A slight increase in operating temperature will increase the efficiency of emission at the expense of the ex-

pected life. At the optimum condition of operation a thoriated-tungsten emitter will give approximately 1 ampere of useful emitted current per square centimeter of surface. Under similar conditions the oxide-coated emitter may give only 0.1 ampere or less of useful current. It might be concluded from this comparison that the thoriated tungsten was more desirable; this would be true if the volume of emitted current per unit area were the important factor. On the other hand, an oxide-coated emitter will require approximately one-tenth as many watts energy per square centimeter for operation as a tungsten cathode. This comparison points to the use of oxide coating for low exciting watts and small emission currents. It should be remembered that only tungsten will withstand positive ion bombardment where very high voltages are involved.

The thoriated-tungsten filament is suitable for operation at fairly high voltages and capable of being reactivated after the emission has been lost owing to temporary overload. The oxide-coated cathode has very high emitting efficiency, but it has a tendency to contaminate adjacent electrodes with small quantities of active emitting material so that emission may take place from these electrodes at relatively low temperatures.

There is no best among the various emitting materials and each has its appropriate place in the field of application, depending on the requirements of operation and the economy of design.

Applications of Emitter Materials. The oxide-coated cathode must always be used for the indirectly heated type because the other cathodes require too high a temperature for satisfactory operation. This application covers nearly all the millions of tubes in use in alternating-current receiving sets. The oxide-coated filament type is best suited wherever quick "heat up" is necessary, and it is essential where a low heating energy drain is necessary because of long hours of operation. This application covers the thousands of tubes in service as telephone repeaters on the toll lines in America.

Thoriated-tungsten and tungsten cathodes find their application in the power tubes used in radio, telephone, and carrier-current transmitters.

In general, it may be said that (1) oxide-coated cathodes are used for tubes up to 100 watts with plate ratings up to 750 volts,* (2) thoriated-tungsten cathodes are used for tube capacities from 100 to

* High plate-voltage ratings up to 20,000 volts have been used in gaseous tubes and large vacuum tubes having sufficient cathode-anode spacing.

1000 watts with plate voltages up to 4000 and 5000 volts, and (3) tungsten is used for cathodes in tubes with ratings from 1 kilowatt up and plate potentials of 5000 volts and higher.

Theories of Work Function. The emission characteristics of cathode materials are more readily understood when viewed in the light of theories of metals suggested in the 1930's.* Under earlier concepts the electrons in metals at ordinary temperatures were assumed to have low energy values. Accordingly, the work function ϕ was assumed to be the total work necessary to remove an electron from a metal. The newer theories advanced by Sommerfeld, Fermi, and others have indicated that the electrons have much higher energies than previously considered. The maximum energy of the electrons in a metal at low temperatures is found to be of the order of 10 volts. This new value has been confirmed by experiments of Davisson and Germer for the reflection of electrons from crystal surfaces. The new findings suggest a simple concept for emission, as shown in Fig. 8.

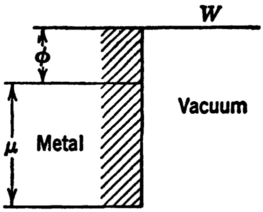


FIG. 8

Here the vertical distances indicate the energies of electrons in electron-volts. Within the metal the electrons may have energies varying from zero to μ. In order to escape into the vacuum outside the surface boundary, a total energy equal to W is required. The difference in energy $W - \mu = \phi$ is ordinarily termed the work function of the material and represents the values given in Table 1.

The concept represented in Fig. 8 for the surface potential boundary is an ideal one, but it fails to account for the effect of the image force which is exerted as soon as the electron emerges. The action of the image charge is shown in Fig. 9 for a small magnified area on the emitting surface.

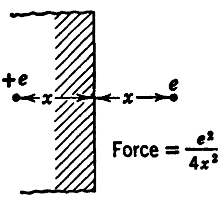


FIG. 9

Whenever an electron of charge e emerges from the surface by a distance x, it is subject to a force due to the image charge $+e$, in accordance with Coulomb's law.

$$\text{Force} = \frac{e^2}{4x^2}$$

* Much of the material in this article has been taken with permission from *Notes on Industrial Electronics* by I. E. Mouromtseff and D. E. Marshall of Westinghouse Electric Corporation.

To move the electron away from the surface against this force requires a varying amount of energy. Since the surface is not smooth, the distance x is hard to define for small values (a few atomic diameters). Hence it is permissible to assume that for small values of x the force will vary linearly, and that for larger values it will vary in accordance with Coulomb's law. This assumption will give a more realistic concept of the energy required, as shown in Fig. 10. The ordinates of the curved boundary line on the right indicate the total energy (ev) that the electron must acquire to move to the given distance from the surface when in free space (zero electric field). Here W again represents the total energy required for emission.

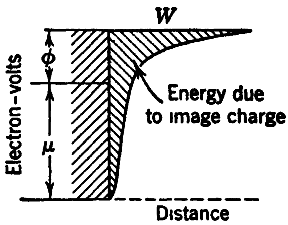


FIG. 10

The emission from hot cathodes does not follow Richardson's equation when an electric field is applied in the surrounding region. Schottky showed that this phenomenon was due to a lowering of the work function or the surface potential barrier by the external electric field. The phenomenon is known as the Schottky effect. Schottky derived the following equation:

$$I = I_0^{e^{3/2}E^{1/2}/KT} \tag{2}$$

where I_0 = current from Richardson's equation with zero field. This equation agrees with experiment to within a few per cent for fields up to 10^6 volts per centimeter for smooth pure metal surfaces. Under normal operating conditions where the applied field is of the order of 2000 volts per centimeter, the increase in emission due to this field is about 10 per cent at 2000 degrees K. A graphical concept of the Schottky effect is given in Fig. 11 which follows from the discussion of Fig. 10. The ordinates of the line labeled "applied field" show the energy given to the electron by the applied field as the electron moves away from the surface. Obviously, this energy must be subtracted (dotted line above) from the energy required for escape in a zero field, giving a net energy shown by the full line. The maximum height of the escape curve is W', leaving a difference $W' - \mu = \phi'$, which is the extra energy or work function for the conditions assumed.

The preceding theory showing the effect of an external applied field can be employed to explain the phenomenon of high-field emission. In Fig. 12, a very high field is shown by the dotted line. When the force of this high field is acting upon an escaping electron, the energy that

must be possessed by the electron is indicated by the full-line energy curve. Since the maximum height of this energy curve is W', it is apparent that those electrons within the surface of the metal that have energies between W' and μ may readily escape even at normal or low temperatures. This is high-field emission. Such emission usually

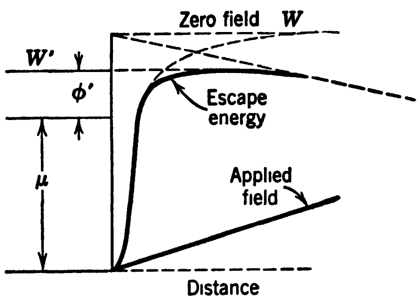


Fig. 11

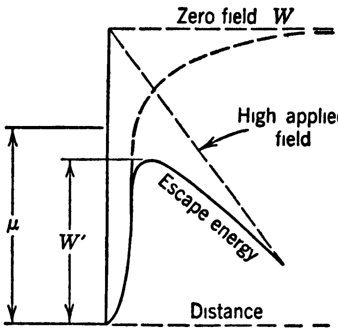


Fig. 12

gives large currents which are likely to be destructive to the cathode surface.

The preceding concepts of work function at surfaces can be applied to other than pure metal cathodes. For the thoriated-tungsten cathode the potential barrier at the surface is pictured in Fig. 13. The ideal

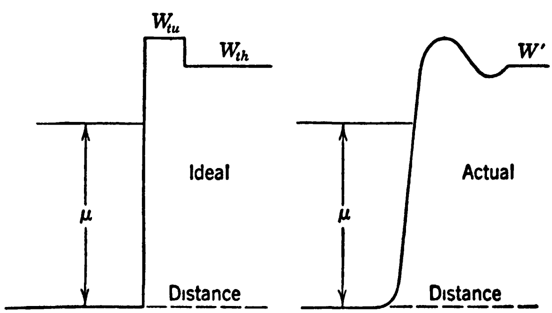


Fig. 13

arrangement on the left shows the higher work function of the tungsten W_{tu} and the lower one of thorium W_{th}. The ideal picture is modified by the image force at the surface, giving the more probable potential barrier relation shown on the right. Apparently the electrons possess sufficient energy to carry them through the hump with enough

remaining energy to escape after emerging into the lower work function region. Another somewhat obscure factor known as polarization enters into the picture. Polarization means that the thorium atoms on the surface of the tungsten reorient themselves in such a manner as to reduce the value of the work function W' below that of thorium on thorium. Note also that an external field will become more effective in reducing the final value of W' after electrons pass over the hump.

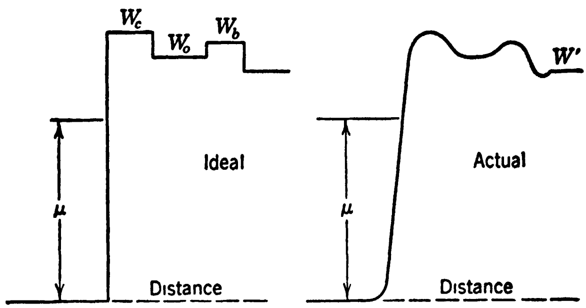


FIG. 14

This explains the fact that the emission from both the thoriated-tungsten and oxide-coated cathodes is much more sensitive to the effect of external fields. Similar diagrams for the required emission energy for oxide-coated cathodes are given in Fig. 14. Here the emitted electrons must pass through two potential barrier humps in order to escape. Electrons in the process of emission must overcome the work function of the core W_c, the oxide W_o, and the surface barium atoms W_b before liberation is achieved.

PROBLEMS

1. What must be the initial velocity of an electron in centimeters per second in order to be emitted from (a) thorium? (b) thoriated tungsten? (c) barium? (d) barium oxide?

2. Assume that a certain tungsten cathode emits 1 milliampere of current at 2000 degrees K. Calculate and plot the curve of its emission through points 1800, 2000, 2200, 2400, 2600, and 2800 degrees K, using Richardson's equation.

3. Repeat Problem 2 for a thoriated-tungsten filament emitting 1 milliampere at 1500 degrees K for a range of 1400 to 2200 degrees K.

4. Repeat Problem 2 for a barium oxide filament emitting 1 milliampere at 800 degrees K for a range of 800 to 1200 degrees K.

5. A certain tungsten cathode is operated by an applied potential of 5.0 volts and 1.0 ampere and supplies a saturation emission current of 50 milli-

amperes. What per cent of the input is transformed into effective energy of emission?

6. At what temperature will a nickel filament emit 5 milliamperes per square centimeter?

7. Using Table 1 and equation 1, calculate the curve of emission for tantalum for temperatures between 2000 and 2800 degrees K.

8. A tungsten filament has a diameter of 0.025 inch and operates at 2500 degrees K. What must be the length of the filament to give a saturation current of 0.5 ampere?

9. A barium oxide-coated cathode consists of a cylinder 0.1 inch in diameter and 1.0 inch long. What will be the saturation current when the cathode operates at 1000 degrees K?

10. Calculate the ratios of the saturation emission currents from tungsten, thoriated-tungsten, and barium oxide-coated cathodes for normal operating temperatures of 2400, 1900, and 1000 degrees K.

REFERENCES

RICHARDSON, O. W., *Emission of Electricity from Hot Bodies*, Second Edition, Longmans, Green & Co., 1921.

LANGMUIR, I., "The Electron Emission from Thoriated-Tungsten Filaments," *Phys. Rev.* **22**, 357 (1923).

DUSHMAN, S., "Thermionic Emission," *Rev. Modern Phys.*, **2**, 381 (1930).

Chapter III

VACUUM DIODES

Unilateral Conductivity. Edison's early experiment in which he observed the "Edison effect" showed that electrons would pass from a hot filament or cathode to a cold plate whenever the plate was at a positive potential with respect to the filament, but that no electrons would pass in either direction if the plate was at a lower potential (negative) with respect to the filament. This property of a device which permits electrons to flow in one direction only is called *unilateral conductivity*. The unilateral conductivity of a tube having two electrodes, one hot and the other cold, was utilized by Fleming for the detection of high-frequency radio waves. This device, known as the Fleming valve, was patented by him in 1905. The Fleming valve was important in the early application of radio telegraphy, and it was one of the leading discoveries in the history of electronics.

Contact Electromotive Force. Volta discovered that, when two different metals were placed in contact and then separated, they acquired electric charges. He also discovered that, when two different metals were placed in an electrolyte and joined by a wire outside the electrolyte, a current was established in the circuit. The action causing such a difference of potential is called contact electromotive force and can be explained on the basis of the electron theory by that intrinsic property of metals known as electron affinity.

To understand this phenomenon, consider the configurations of the blocks of tungsten and nickel shown in Fig. 1. If the two blocks are placed as in part *a* and are in an uncharged state, there will be no difference of potential between them. Now, if they are brought together as in *b*, their contact surfaces must come to the same potential, but, when they are separated again to position *a*, they will exhibit a difference of potential and an electric field will exist between them. Joining the two blocks by a conductor as in *c* accomplishes the same result as moving them to position *b* and back to *a*. After the conductor connects the blocks in position *c*, a difference of potential exists between the blocks. The differences of potential mentioned here must be

measured with apparatus that does not require current for its operation.

The reason for this strange phenomenon is as follows: The surface of the tungsten block has a potential energy barrier equal to its work function of 4.52 volts. Similarly, the nickel block has a potential energy barrier equivalent to 4.0 volts. When the two blocks are brought in contact as in *b*, the potential energy barriers must come to a common level. Since the nickel block has the lower electron affinity, electrons move from it across the boundary of contact into the bounding surface of tungsten and are held there by the image forces. When the blocks are separated the image charges are too far apart to have appreciable effect, the transferred electrons redistribute themselves over the blocks, and the nickel block is positive with respect to the tungsten. The difference of potential now existing is equal to the difference in the work functions, or $4.52 - 4.0 = 0.52$ volt, and is the *contact electromotive force* for these metals. In the action described the metal with the lower work function becomes positive, and hence this metal is said to be electropositive with respect to the other.

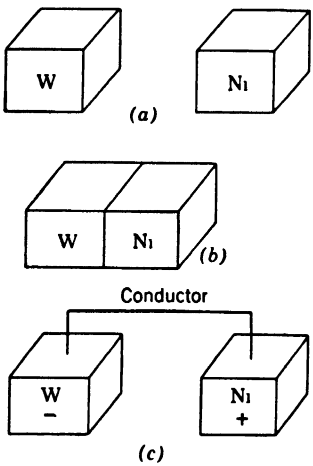


FIG. 1. Configurations illustrating contact electromotive force.

In the manufacture of electron tubes the electrodes are frequently made of different metals. These electrodes are usually joined by a circuit (conductor) similar to configuration *c* in Fig. 1. The contact emf's between these electrodes are added algebraically to any other differences of potential in these circuits. While these contact emf's are small in magnitude (usually a fraction of a volt), they may have a bearing on the operation of the tubes in which they exist when relatively low potentials exist on the electrodes. In high-mu triodes (such as 6Q7), which have a mu of approximately 100, the contact emf between the cathode and grid may be sufficient to provide the negative grid bias.*

Theory of Rectification in Vacuum. The theory of unilateral conductivity is simple and follows directly from the principles covered in

* Grid bias will be discussed in Chapter IV.

previous chapters. It will be assumed first that a hot electrode (cathode) and a cold electrode (anode) are placed in a tube where a good vacuum exists. The anode will be free (not connected to anything), as shown in Fig. 2. The battery A supplies current for heating the cathode, and when the optimum temperature is reached a cloud of electrons will surround the cathode. Each of the individual electrons in the cloud is thrown off with some initial velocity, and it moves out toward the anode against an attraction from its image positive charge on the filament. Since the anode is of neutral potential, it does not influence the electron. In all probability the kinetic energy of the electron due to initial velocity of emission will be overcome before it reaches the anode, and it will drop back to the filament. Now as point x is moved from a towards b, the current in the filament rises, which raises the temperature of the cathode, increases the rate of emission, and increases the initial velocity of emission of electrons. As this process is continued, a point will be reached where a few electrons will have a sufficient initial velocity to carry them over to the anode where they will "stick." The presence of these electrons on the anode will give it a negative charge, and this charge will now repel those electrons that approach the anode. Obviously, a point of equilibrium will soon be reached where no more electrons will become attached to the anode for the given temperature of the filament. If, now, point x is moved nearer to b, the initial velocity of some electrons will be raised so that they can overcome the small repulsion of the anode and land on it. Again, the negative potential on the anode will rise and a new condition of equilibrium will be reached. Since the anode is free and insulated electrically, the electrons that land on it have no avenue of escape.

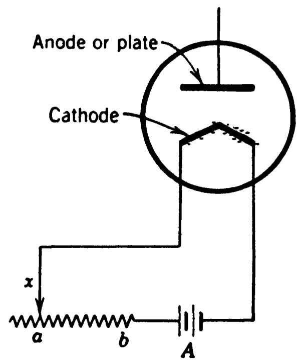


FIG. 2

If the circuit of Fig. 2 is changed by connecting the anode electrically to the filament as shown in Fig. 3, the action of the device is altered considerably.

The joining of the electrodes introduces a contact emf if different metals are involved. Assuming that the filament of the cathode is tungsten and the anode is nickel, a contact emf of 0.52 volt will exist with the anode positive with respect to the cathode. When the temperature of the cathode becomes sufficiently high, the initial velocity of

emission of some electrons will carry them to the anode as in the preceding case. This movement of electrons to the anode will be affected by the contact emf now present, and for the metals assumed the positive value on the anode will aid (attract) the passage of the electrons. Thus the electron transfer will be larger than in the preceding case. The electrons that land on the anode will not accumulate as before, since they are free to return to the cathode via the external conductor. Thus a very small continuous electron current will be produced by the

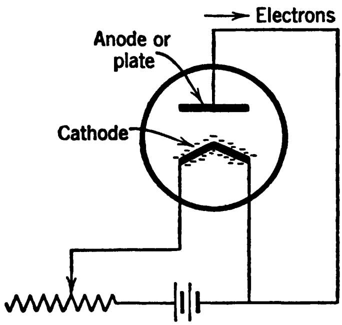


FIG. 3

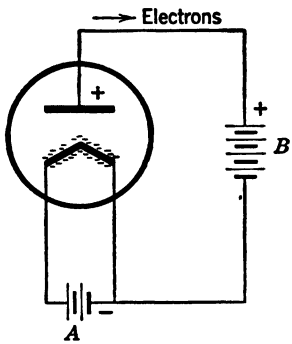


FIG 4

electrons passing through the vacuum from cathode to anode and returning via the external path.

The addition of a battery in series with the anode circuit, as shown in Fig. 4, gives one common circuit for the use of the two-electrode tube. The theory of action becomes an extension of the preceding discussion. The hot cathode emits a cloud of electrons. Each electron that leaves the cathode is subject to a number of forces and factors that determine whether it moves over to the anode. The first factor is the initial velocity of emission determined by the temperature of the cathode. A second factor is the positive image charge on the cathode. A third and new factor is the attraction of the positive charge on the anode for the electron. A fourth factor is the negative space charge, which will be explained in the next article. The contact emf, if present, is a minor factor. The initial velocity of emission and the attraction of the anode are positive factors tending to carry the electron over to the anode. The attraction of the cathode is a negative factor and the negative space charge in general is also negative or opposing in its action. If the polarity of the B battery in Fig. 4 is reversed so that the anode becomes negative with respect to the filament, the anode will repel the electrons that approach it and, except for a very low negative potential

(fraction of a volt), no electrons will ever be able to land on the anode. Here the negative anode becomes the controlling factor and accounts for the unilateral conductivity of this type of vacuum tube. It should be noted that the positive potential placed on the hot cathode will not secure electrons from the negative anode because the anode is cold and is not a source of electron emission.

Negative Space Charge. The individual electrons emitted from a cathode are negative charges and as such they exert a force of attraction or repulsion upon surrounding charged bodies and particles. The cloud of electrons emitted from a filament thus acts as a charge or field in the region surrounding the cathode (Fig. 5). This cloud is very dense close to the filament and grows thinner rapidly as the distance from the filament is increased. The action of this space charge on an individually emitted electron is easy to analyze. As the electron breaks through the surface of the filament, it is repelled by the millions of other emitted electrons lying close to the filament as well as by those farther away. Thus at this point the space charge exerts a very powerful opposing effect upon this electron. As the electron moves out farther, as illustrated in Fig. 5, it is repelled toward the filament by all electrons in space between it and the anode and is repelled (aided) toward the anode by all electrons in space between it and the cathode. Obviously, the effect of space charge varies with the position of the electron under consideration. It is retarding for positions close to the cathode and aiding at positions nearer the anode.

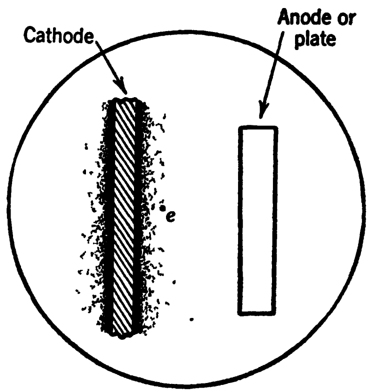


FIG 5 Negative space charge surrounding a hot cathode.

The influence of negative space charge may be explained by means of potential distribution curves. Figure 6 (part a) illustrates a two-electrode tube having electrodes C and A with a difference of potential of 100 volts between them. Obviously, the potential at C is zero or ground potential, and, as the point of view moves from C to A, the potential must rise from zero to 100 volts. The manner of rise of potential from C to A may be shown graphically by curves as in part b of Fig. 6. If C and A consist of two large cold parallel plates, the change of potential along a line near the center of the plates will be uni-

form (a constant potential gradient) and may be represented by the straight line *a*. If *C* is a small round wire (cold) and *A* is a hollow concentric cylinder surrounding *C*, the potential gradient near the wire will be high (owing to strong electrostatic field) and then will fall off as *A* is approached. Such a gradient may be represented by the dotted curve *b*. If *C* is a hot filament wire, the potential distribution may be

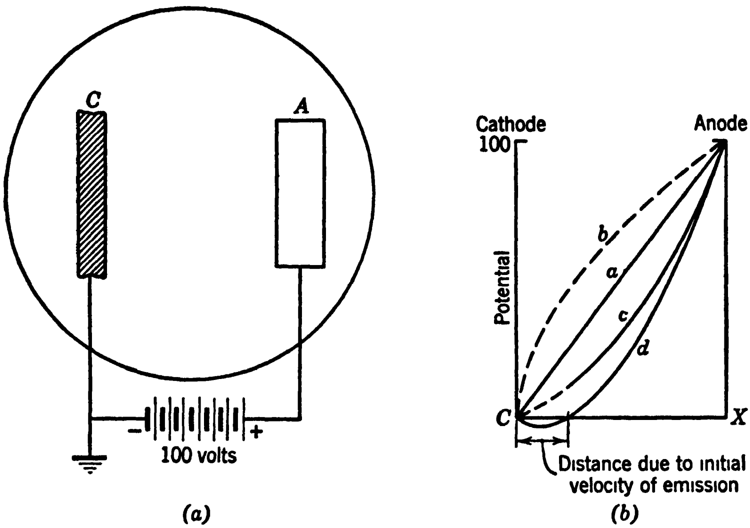


FIG. 6. Potential distribution in a two-electrode vacuum tube.

a curve like *c*. Here the negative space charge surrounding the hot cathode depresses the potential in that region so that the curve dips below a straight line near *C*. If the cathode is heated to a high value, the electron emission and the space charge around *C* may become great enough to cause the potential in space near *C* to be actually lower than that of *C* and to give a potential distribution indicated by line *d*. If one imagines an electron being carried the distance from *C* to *X*, Fig. 6, curve *d*, it is easy to understand the depressing or opposing influence of the space charge near the cathode.

The negative space charge is very effective in limiting the number of electrons that pass to the anode of tubes having a high vacuum. It is so effective that it was once believed that there would not be any electron current in a perfect vacuum. Space charge can be overcome by high anode potentials and its effect can be neutralized by the presence of positive ions, as will be explained later.

Construction of Anodes. The anode is always the *electron-collecting electrode* in an electron tube. In the first tubes built, the anode was constructed in the form of a little plate or sometimes two little plates connected in parallel and placed on each side of the filament type of cathode. The name plate has persisted down to the present time for the anodes of tubes used in communication circuits. For tubes used in power circuits and photoelectric devices, the term anode is more generally used. It is unfortunate that the term plate persists as applied to the anode because plate is a more fitting term for the electrostatic deflection electrodes used in cathode-ray tubes and similar devices. In describing tubes in this textbook, the author will attempt to use whichever term, *anode* or *plate,* has the most general acceptance in handbooks giving tube characteristics.

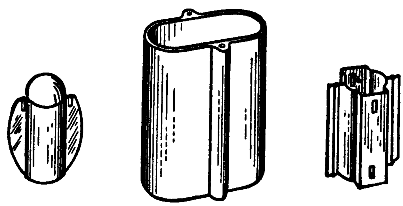


FIG. 7. Typical plate structures. (Courtesy Radio Corporation of America.)

The anode or plate is usually constructed so that it surrounds the cathode in the vacuum type of electron tube. This construction reduces the length of electron path and the resultant potential drop from cathode to plate. The typical plate structure is a hollow tube of circular, oval, or rectangular cross section, as illustrated in Fig. 7.

Many different materials have been used for constructing anodes. The desirable mechanical properties for this service are easy workability, mechanical strength at high temperatures, high thermal radiation, and low vapor pressure. Tungsten has desirable properties, but it is hard to form and hence has little use today for anode construction. Molybdenum has all desired properties except high heat emissivity. This weakness is overcome by use of radiating fins and a roughening of its surface and coating with materials such as zirconium powder (sintered).

Graphite has considerable use as an anode material. Its principal weakness is that it absorbs more gas than other materials. Nickel has a low melting point but can be used for low-power tubes. It is usually carbonized by a process wherein a well-adhering layer of amorphous carbon is deposited on the nickel. This process provides a heat emissivity approaching that of a black body and serves to reduce secondary emission. Tantalum is a satisfactory material for anode construction and is finding an increasing application in this field. Like molybdenum,

it usually requires fins on the anode and a roughened surface to increase the heat-radiating area.

The principal limitation in the design of an anode is its heat-radiating ability. The anode receives heat from two principal sources: (1) the heat radiated from the cathode and (2) the heat generated by the impact of the impinging electrons from the cathode.

If an electron is not intercepted in its flight from the cathode to anode, it strikes the anode with a kinetic energy equal to $\frac{1}{2}mv^2$. In high-vacuum tubes this kinetic energy is equal to the anode potential times the charge on the electron, or

$$e_b e = \tfrac{1}{2}mv^2$$

All this energy must be transformed into heat. Thus the power represented by anode voltage times the anode current is transformed into heat energy at the anode, and this energy must be dissipated by radiation, convection, and conduction to the outside medium. It is of interest to know that the bombardment of the anode by the electrons does eject numerous electrons from the anode through secondary emission. The electrons of secondary emission fall back into the anode and do not affect the magnitude of the anode current in the diode.

Characteristics of a Vacuum Diode. A two-electrode tube having a gas pressure of about 10^{-8} atmosphere or less is called a high-vacuum diode. A vacuum diode has two characteristics that are fundamental in understanding the operation of two-electrode tubes and the multi-electrode tubes to be covered in succeeding chapters. The two variables in the diode are the electron emission of the cathode and the potential applied between the cathode and the plate. If one is held constant and the other is changed, the two characteristics can be observed.

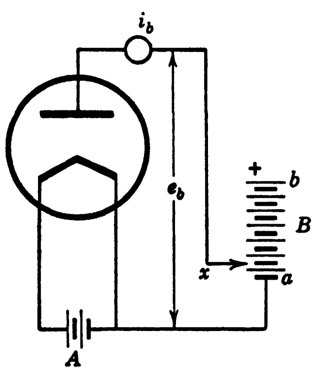


FIG. 8. Circuit for determining the voltage saturation characteristic of a vacuum diode.

The plate-voltage characteristic of a diode can be determined by connecting the tube in the circuit of Fig. 8. The filament or heater current should be held at rated value to give normal electron emission, and then the potential between the plate and cathode e_b can be varied from zero up to a high value. The resulting change of plate cur-

rent i_b with the variation in plate voltage is given in Fig. 9 for two different values of filament voltage. In both, the lower part of the curve rises rather slowly at first, then more rapidly for a time, and ultimately bends over and tends to flatten out or become horizontal. The slower rise of the plate current initially is caused by the strong retarding influence of negative space charge. For the lower curve the plate current begins to flatten out after 60 volts is reached. At this point the

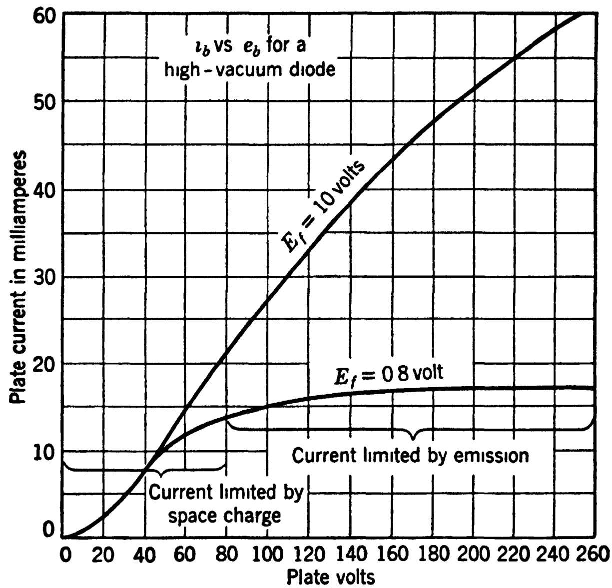


FIG. 9. Plate-voltage characteristic of a vacuum diode.

plate is attracting the majority of all electrons emitted, and a kind of "saturation" is rapidly approached as the plate voltage is raised further. If the filament temperature and resulting emission are raised by impressing a higher voltage on the filament (upper curve), saturation will be reached at much higher plate potential values. This is to be expected, for, with the emission of electrons greatly increased, the space-charge effect will be increased and also a higher plate voltage will be needed to attract the additional electrons before saturation is reached. The characteristic curves of a vacuum diode under discussion are frequently called voltage saturation curves.

The plate-current plate-voltage (i_be_b) characteristic curve for a vacuum diode is represented by an equation due to Childs.

$$i_b = A \frac{e_b^{3/2}}{x^2} \tag{1}$$

where i_b = total plate current.
e_b = total plate voltage.
x = distance between electrodes.
A = constant which depends on the geometry of the electrodes.

For a given tube where x is constant, the equation reduces to

$$i_b = ke_b^{3/2} \tag{2}$$

It should be noted that Childs' equation holds where all points of the cathode are at the same potential and at the same distance from the plate as in the heater type of cathode. This condition does not exist where the cathode is a filament having a varying potential along its length. Childs' equation also does not hold after saturation is approached or if electrons are emitted with an initial velocity.

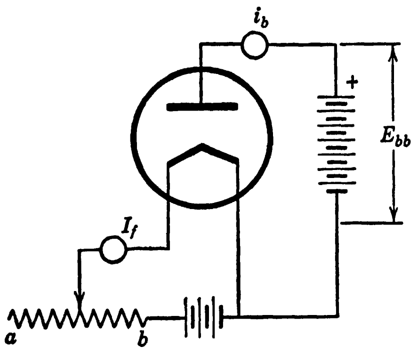


FIG. 10. Circuit for determining the temperature saturation characteristic of a vacuum diode.

The current covered by Childs' equation and the preceding discussion should never be confused with emission current. After saturation the plate current does represent the total emission current but not before.

A second characteristic of the diode is determined from the circuit of Fig. 10 by holding the plate supply potential E_{bb} constant and varying the filament current from zero up to the maximum permissible.† The variation of plate current for a series of plate supply potentials is illustrated in Fig. 11. The trend of these curves is approximately the same as those of Fig. 9 but the cause of the trends is reversed. The initial plate current is limited by emission, since the plate is attracting all the electrons being emitted. The trend of this part of the curve is similar to the curve that follows Richardson's equation. The i_b curve

* The exponent of e_b is 3/2 for parallel-plane diodes. Other configurations require different exponents.

† See symbols for electron-tube circuits in front of book.

tends to flatten out at the higher filament currents because of the negative space-charge effect. Thus at low filament currents and temperatures the space charge is negligible and the plate attracts all the electrons emitted, but as the cathode emission rises the space-charge effect becomes stronger and stronger and finally prevents the plate

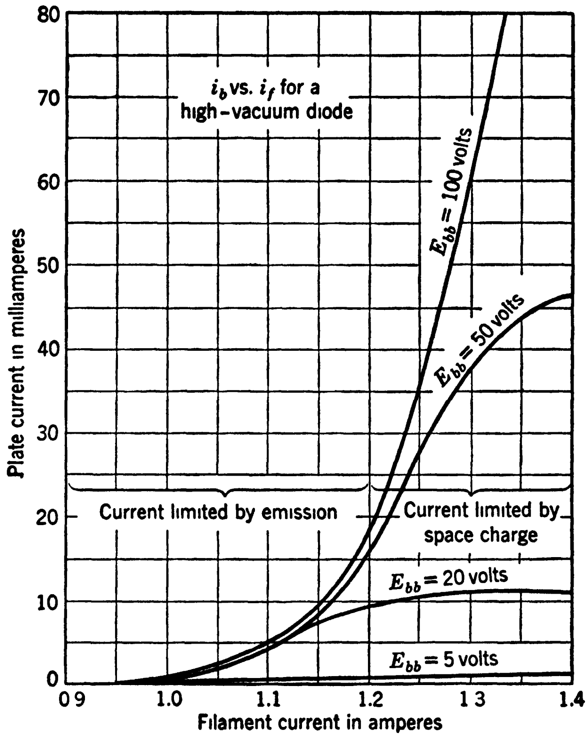


FIG. 11. Cathode-current characteristic of a vacuum diode.

from attracting additional electrons even though more are available in the expanding electron cloud. The curves of Fig. 11 are sometimes called temperature saturation curves. In these the plate attraction is constant with the negative space charge growing stronger, whereas in Fig. 9 the negative space-charge effect is constant and the plate attraction is growing stronger. In Fig. 11, the plate-current curve for $E_{bb} = 100$ volts rises more rapidly from zero because of the Schottky effect (page 35).

Rectification of Alternating Currents. The unilateral property of the two-electrode tube suggests a simple method of rectifying alternat-

ing current or transforming it to a unidirectional current. A two-electrode tube of the vacuum type is connected in a circuit as indicated in Fig. 12 with an alternating voltage impressed between the cathode and plate and a low-voltage a-c supply for the cathode heater. On the top half of the alternating-voltage cycle (part b) the plate is positive and electrons will pass to it. The electron flow will produce a current loop as indicated in part c. When the voltage drops to zero the current drops to zero, and when the plate becomes negative it repels the electrons emitted and no current results. Thus the alternating voltage results in a pulsating direct current which is termed half-wave rectification. If it is desired to utilize the idle portions of the alternating-voltage loops, a second rectifying tube may be added as shown in Fig. 13. An analysis of the operation of this circuit will show that, when the alternating-current supply main L_1 on the left is positive, the electrons move in the direction shown by the full-line arrows. Conversely, when the supply main L_2 on the right is positive, the electrons move as indicated by the dotted-line arrows. This action gives a full-wave rectification of current in the d-c load circuit as illustrated by part c of Fig. 13. This pulsating form of direct current is satisfactory for charging storage batteries and for electroplating. Whenever a steady d-c voltage or current is necessary, the pulsations can be smoothed out by filters as will be explained in Chapter XII on rectifiers.

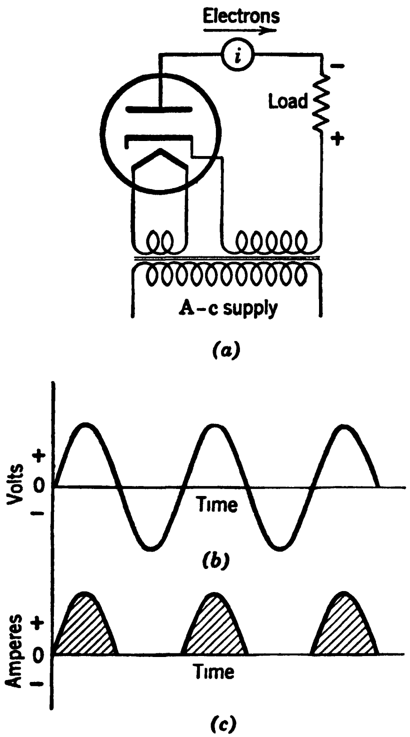


FIG. 12. Circuit and curves showing half-wave rectification of alternating current.

A simple circuit of a rectifier, a filter, and a potential divider is given in Fig. 14. This complete unit is called a power pack and is widely used in radio receivers, radio transmitters, and other electronic assemblies as the source of d-c voltage. The resistor at the right of

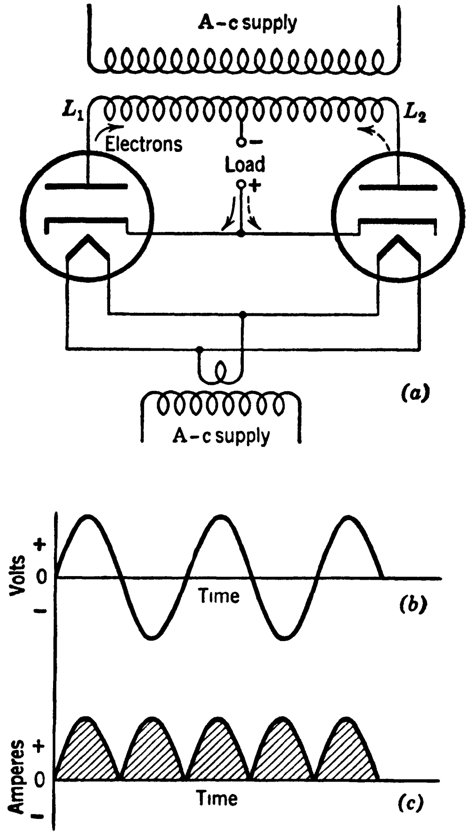


FIG. 13. Circuit and curves showing full-wave rectification of alternating current.

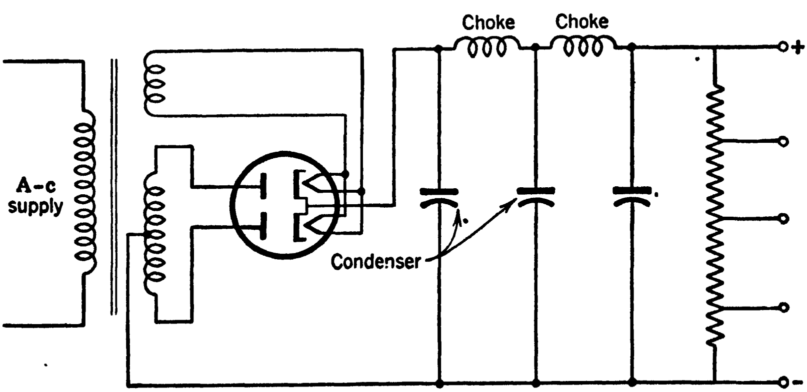
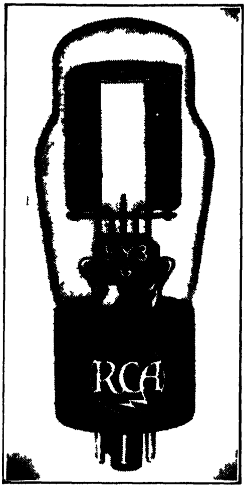
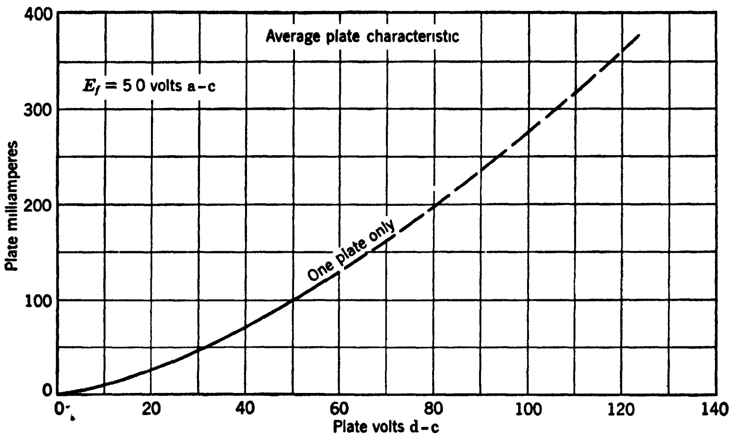


FIG. 14. Circuit for a power pack.



FULL-WAVE HIGH-VACUUM RECTIFIER

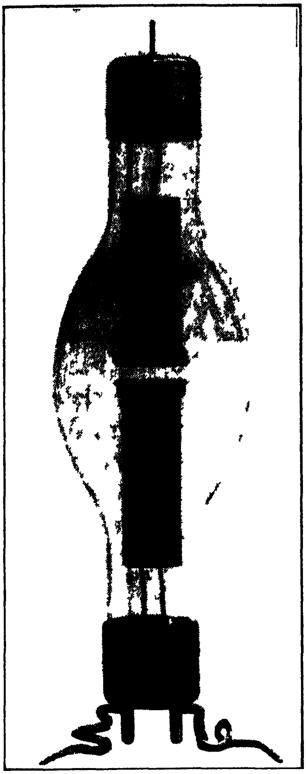
Filament—coated	
Voltage	5 0 a-c volts
Current	2 0 amp
Peak inverse voltage, max	1400 volts
Peak plate current per plate, max	375 ma
With Condenser-Input Filter	
A-c plate voltage per plate (rms), max	350 volts
Total effective plate-supply impedance per plate, min	50 ohms
D-c output current, max	125 ma
With Choke-Input Filter	
A-c plate voltage per plate (rms), max	500 volts
Input-choke inductance, min	5 henries
D-c output current, max	125 ma



TWIN DIODE

Heater—coated unipotential cathodes	
Voltage	6 3 volts
Current	0 3 amp
Direct interelectrode capacitances	
Plate 1 to cathode 1	3 0 $\mu\mu f$
Plate 2 to cathode 2	3 4 $\mu\mu f$
Plate 1 to plate 2, max	0 10 $\mu\mu f$
Rectifier	
A-c plate voltage, max	117 0 volts
D-c current, max	4 0 ma

Fig. 15. Vacuum rectifier tubes and characteristics. (Courtesy Radio Corporation of America.)



KENOTRON
WL-456

HIGH-VOLTAGE VACUUM RECTIFIER

General Characteristics	
Air-cooled diode	
Filament voltage	11 volts
Filament current	20 amp
Filament heating time	30 sec
Net weight	2 lb
Shipping weight	6 lb
Maximum Ratings	
Anode voltage, peak inverse	140,000 volts
Anode current, peak	500 ma
Anode current, average	60 ma

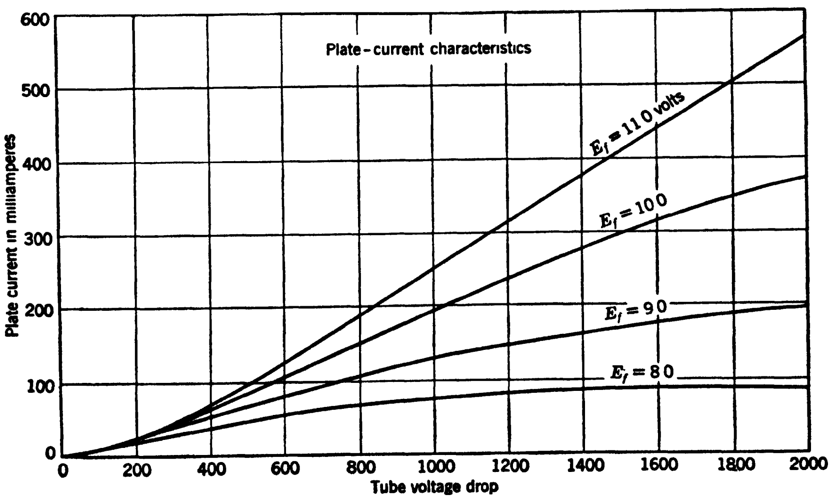


FIG. 16. Rating and characteristics of a high-voltage vacuum diode (Courtesy Westinghouse Electric Corporation.)

the circuit with its numerous taps serves as a potential divider for furnishing a selection of voltages.

Applications of High-Vacuum Diodes. The applications of the vacuum diode are determined by its inherent properties. It is essentially a rectifier. Its plate current is limited by the negative space charge, so it has a small current rating. The potential drop from cathode to plate is rather high and varies with the plate current. This large voltage drop results in a low efficiency of rectification unless high voltages are used. In the low-voltage field the vacuum diode is used as a detector of radio signals and in power packs for radio receivers requiring currents of low magnitude. The characteristics of two tubes for this type of service are illustrated in Fig. 15. The vacuum diode will withstand a high inverse voltage, i.e., a high negative potential on the plate, without an arc-back. This makes it very desirable for high-voltage, low-current rectification. Thus it is widely used for furnishing direct current of the order of 1000 to 100,000 volts for X-ray machines, radio transmitting tubes, smoke precipitators, and dust eliminators. Vacuum diodes for the latter class of service are known as *kenotrons*. A typical kenotron and its rating and characteristics are illustrated in Fig. 16.

Problems

1. A certain vacuum diode has an emission of 10 milliamperes for a plate potential of 50 volts. Assuming that Childs' law holds for the tube, calculate and plot the saturation-emission current for 20-volt steps from zero to 100 volts on the plate.

2. Assume in Fig. 16 a cathode-anode separation of 2 centimeters and a peak applied potential of 20,000 volts What will be the velocity of an electron when it hits the anode if it left the cathode with zero initial velocity? What will be the energy of this electron on impact in (a) electron volts? (b) joules?

3. What is the average internal or plate resistance of the upper vacuum diode shown in Fig. 15?

4. Calculate Problem 3 for the kenotron of Fig. 16, when $E_f = 11$ volts.

5. The kenotron of Fig. 16 is employed for half-wave rectification, feeding a load resistor of 10,000 ohms. What is the voltage across the load resistor when the drop across the tube is 1000 volts? What power is being dissipated at this instant in the load? within the tube (including filament power)?

REFERENCE

McArthur, E. D., "Electronics and Electric Tubes," *Gen. Elec. Rev.*, December 1933.

Chapter IV

GRID-CONTROLLED VACUUM TUBES

Three-Electrode Vacuum Tube. In 1907 DeForest added a third electrode to the Fleming valve and called the new device the audion. The third electrode, which he called a grid, consisted of a zigzag wire which was placed between a heated filament and a plate (see Frontispiece). The value of the grid lay in its ability to control the electron current between the cathode and the plate, and this invention constituted the most important development of the twentieth century. It has extended the field of communication by wire across continents, it has made possible radio communication around the world, and it is now revolutionizing the use and control of electric power in the industrial world.

The modern three-electrode vacuum tube, known by the family name of triode, uses cathodes and plates like those described in the preceding chapters. The grid is usually a coil of fine wire wound in the form of a helix and interposed between the cathode and plate. The construction may consist of circular concentric cylinders as shown schematically in part *a* of Fig. 1, or it may utilize oval or oblong cylinders as in part *c* of the same figure. The standard symbols for the triode are given in part *b* of Fig. 1; the upper symbol represents the filament type of cathode and the lower the heater type. The cathode heater is not an electrode.

Theory of Grid Action. The grid of the three-electrode vacuum tube functions by a change of the charge residing upon the grid. This change in charge (and potential) serves to control the electron stream between the cathode and the plate. The process of control can be visualized in a number of different ways. One simple visualization follows from Fig. 2, which is a schematic diagram of a cross section taken through the axis of a three-electrode tube. Here the helical wire grid appears as circles. When the cathode is heated it will emit a cloud of electrons, part of which will be attracted over to the plate by its positive charge. The electrons that go to the plate must pass through the meshes of the grid and hence will be affected by the potential residing

on the grid. Assuming that the grid is positively charged with respect to the cathode, an electron in position *a* will be subject to several forces acting on it: first, the initial velocity of emission; second, the

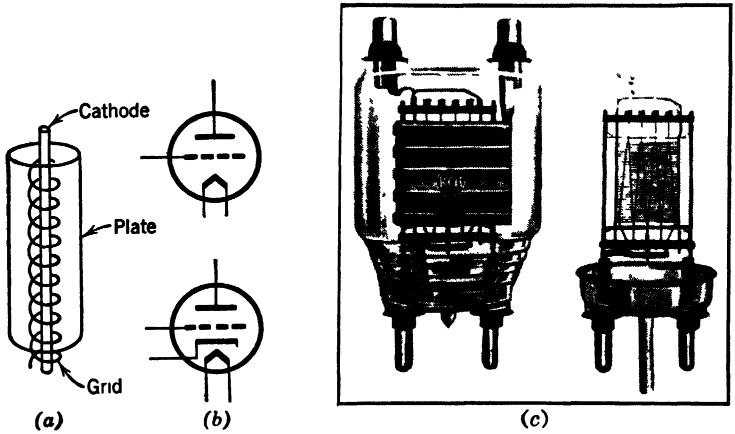


FIG. 1. (*a*) Schematic triode. (*b*) Triode symbols. (*c*) Triode parts and assembly. (Courtesy of Radio Corporation of America.)

attraction back to the cathode due to the image positive charge left on the cathode; third, the influence of negative space charge; fourth, the attraction due to the plate; and, last, the attraction of the charge on the grid wires. The attraction of the grid wires will be in the directions indicated, but the resultant of all these forces will be toward the plate. Thus the grid aids (controls) the passage of electrons to the plate. If an electron progresses to the point shown at *b*, it will be subject primarily to three influences: (1) its instantaneous velocity, (2) the attraction of the plate, and (3) the resultant attraction of the grid wires. This resultant force of the grid wires is zero, and hence the grid is now ineffective, but it has accomplished its task in helping the electron to escape from the cathode and the repulsion of negative space charge. An electron in the position *c* is subject to the same group of forces as any other electron, but the resultant of the attraction of the grid wires and the pull of the plate potential is directly toward the grid wire, and hence this electron will land on the grid. Electrons moving on a line

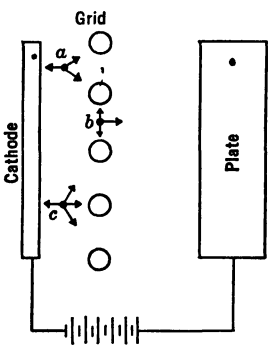


FIG. 2. Forces acting upon electrons in a triode.

passing to one side of the grid wires will probably pass to the plate because of their high velocity and because the grid potentials are usually low relative to the plate potential.

If the potential of the grid is now made negative with respect to the

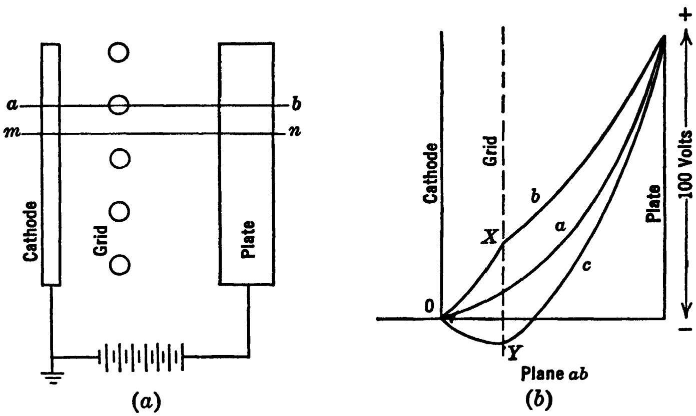


FIG. 3. Potential distribution in a triode (plane through grid).

cathode and a like process of reasoning is applied, it is evident that the action of the grid will always oppose the passage of the electrons to the plate. The degree of opposition offered will depend on the magnitude of the negative potential. It is apparent that a strong negative potential or negative bias may bar all electrons from passing to the plate or from passing to the grid itself. If the grid is given a small negative bias, changes of potential superimposed on this bias will produce corresponding swings in the cathode-plate current.

A second method of analyzing grid action involves a study of the potential distribution in the three-electrode tube. If in Fig. 3, part *a*, a plane *ab* is passed perpendicularly to the axis of the tube, it will show a potential distribution from cathode to plate as in part *b*. Let it be assumed that the cathode is at zero or ground potential and that the plate is maintained at a potential of 100 volts above the cathode by the battery. If the grid were omitted and the cathode heated, the potential distribution would be given by curve *a* of Fig. 3*b*, the depression of the curve being due to the negative space charge. Now, if the grid is added and raised to a small positive potential, the potential distribution will be raised to that of curve *b* with point *X* at the exact potential of the grid. Again, if a negative potential is applied to the grid, the potential distribution will fall to that of curve *c*, where *Y* is the negative potential applied to the grid. The change in the potential

distribution will not be so pronounced for other planes in the three-electrode tube. Thus if a plane is passed through *mn* of Fig. 3, part *a*, the potential distribution will be represented by curves *a′*, *b′*, and *c′* of Fig. 4. Planes passed through intermediate points would show potential distribution curves varying between the limiting cases illustrated in Figs. 3 and 4. In any case the effect of the potential on the grid is obvious. When the grid potential is raised, it aids the electrons in breaking away from the cathode and moving to the plate, whereas a negative potential opposes the electron movement to the plate.

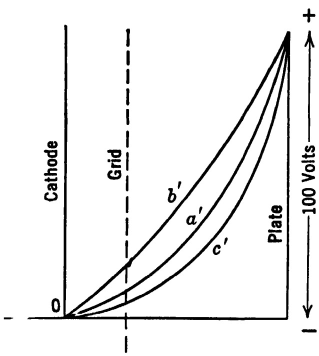


FIG. 4. Potential distribution in a triode (plane between grid wires).

It is evident from the preceding paragraph that the closeness of the wires in the grid will affect the degree of the grid control of the electron stream. If a coarse mesh is used, electrons will pass midway between the wires with a relatively small influence from the grid. A fine or close mesh will give a more uniform and effective control. Some tubes have a coarse grid near the center and a fine mesh at the top and bottom. These are known as variable-mu tubes and their advantage will be pointed out later.

A third method of visualizing grid action in a triode is through a study of patterns in the electric field in the cathode-plate space under a variation of the potential applied to the grid. Pictures of possible electric fields are given in Fig. 5. Part *a* of this figure shows a field of uniform direction and distribution which would exist without any grid (diode). In part *b*, it is assumed that a grid has been inserted and its potential has been made the same as the potential of the space it occupies. Under these conditions the grid does not influence the field. In part *c*, the grid has been connected to the cathode so that it is at zero or cathode potential. Since the grid is closer to the plate than the cathode is, many of the lines of the electrostatic field now emanate from grid to plate. In part *d*, the grid has been made slightly positive with respect to the cathode. Here the field lines from the negative cathode lead to both grid and plate. This appears the same as in part *b* where the grid was positive, although there may be a difference in the magnitude of the field in the cathode-grid space. A change of the grid potential to negative in part *e* of Fig. 5 causes a reversal in the direction of some lines in the electric field. Now, some of the field

lines from grid lead to the cathode. The pictures of the electric field are completed by part *f* where the grid is made strongly negative so that all field lines to the cathode have been reversed. An analysis of the action of these various patterns of electric-field direction and distribution upon electrons released at the cathode will show the effectiveness of grid control.

The early triodes and most of those in use today have the grid located between the cathode and the plate. It is possible to place the grid

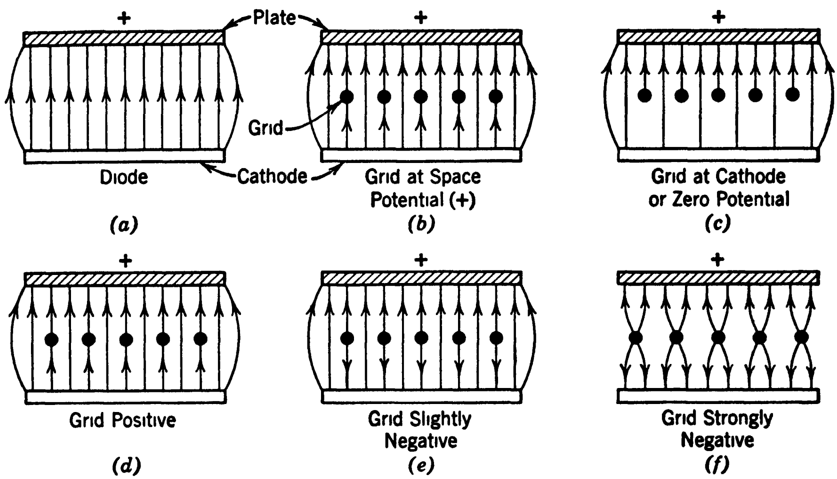


FIG. 5. Electric field distribution in triodes under different grid potentials.

in other positions. Thus one type of tube has a central rod for a grid, which is surrounded by a helical coil (cathode) with the plate outside the cathode. This is known as an internal-grid tube. The control of the grid is not so effective, but the closeness of the cathode to the plate results in a relatively large cathode-plate current, with the expenditure of little or no power in the grid circuit. This tube is useful for certain types of relay work, oscillograph amplifiers, and speech-frequency amplifiers.

A conventional circuit for the three-electrode tube using batteries for the necessary applied potentials is shown in Fig. 6. The filament type of cathode is heated by a battery called the *A* battery, the plate secures its source of potential from another battery known as a *B* battery, while the input or potential variation to the grid is connected between the grid and the cathode. If the grid is to be kept normally negative, a battery called a *C* or biasing battery is connected into the grid supply circuit. The voltages supplied by these three batteries are

represented by the symbols E_{ff}, E_{bb}, and E_{cc} as shown in Fig. 6. If at any time the grid becomes positive with respect to the cathode, it will attract a few electrons to itself, and these electrons will return to the cathode via the grid-cathode circuit. These electrons constitute the grid current. The plate current is usually several times as large as the grid current when the grid is slightly positive. When the grid is nega-

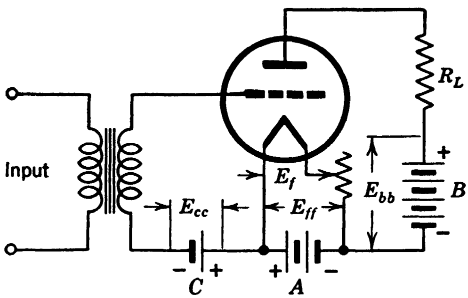


FIG 6 Conventional circuit for a triode

tive, it is generally assumed that the grid circuit does not carry any current, although this is not rigidly true as will be shown later.

A *free grid* is one that is isolated from any circuit through which a continuous stream of electrons may pass. This may mean that the grid is entirely free as in part *a* of Fig 7, or it may be connected to a series circuit containing a condenser which will block any continuous

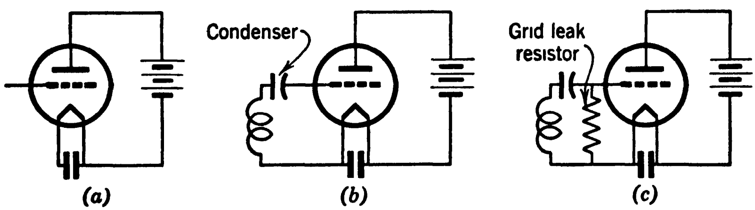


FIG. 7 Free-grid and grid-leak circuits

flow of electrons in either direction (part *b*, Fig. 7). Before a triode containing a free grid is energized, the grid will probably be at zero or ground potential. As soon as the cathode and plate are energized (part *a* or *b* of Fig 7), electrons from the cathode may move out under their initial velocity of emission and attraction of the plate and land on the grid. Since these electrons cannot leave the grid through any external conductor, the grid will assume a negative potential of such a magnitude as to bar any more electrons from landing on it. Also, if a

varying signal voltage is impressed on the grid through a condenser, it will provide transient changes in grid bias which may permit the grid to go even more negative until it serves as a complete block to any electron movement to the plate. The accumulation of a strong negative charge on a free grid makes it impossible to use the triode in any useful circuit. Accordingly, it is necessary to provide some path for the accumulating charge to leak back to the cathode. In some circuits the input circuit itself will provide a natural leak for the grid. In other circuits where the input to the grid contains a condenser as in part *b* of Fig. 7, it is necessary to provide a high-resistance path called a *grid leak* as illustrated in part *c* of Fig. 7. A grid leak has a value from one-half up to several megohms.

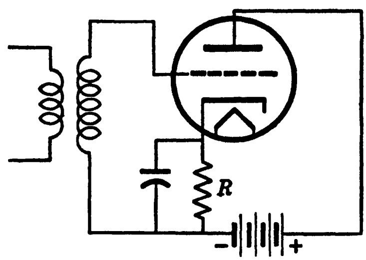


FIG. 8. Circuit for producing negative grid bias.

The triode and the multielectrode tubes to be described later are operated with the grid held negative with respect to the cathode. This negative grid or bias prevents any normal electron current from cathode to grid which would represent a useless current and power loss, and it also provides better operating characteristics in the tubes themselves. The grid bias may be provided by three methods. One method using a *C* battery is illustrated in Fig. 6. A second method provides the desired bias by the selection of a suitable value for the grid leak shown in part *c* of Fig. 7. The leak of electrons back to the cathode through this high resistance provides an *RI* drop with the grid held negative. A third method for obtaining a negative bias utilizes a voltage drop arising from the direct current in the cathode-plate circuit. This is illustrated in Fig. 8 where the current through *R* provides the desired potential drop. The condenser in parallel with *R* provides a low-impedance shunt path for any signal frequencies in both the grid and plate-cathode circuit. In a few applications a low grid bias is provided by the contact emf between the cathode and the grid.

Characteristics of a Three-Electrode Tube. The important characteristics of the three-electrode tube shows how the anode or plate current changes when either the grid voltage or the plate voltage is varied, with the other held constant. The transfer characteristic of a triode may be determined through the use of the circuit given in Fig. 9. The cathode is heated to its normal operating temperature and a

voltage E_{bb} is applied to the plate. The grid is supplied by a circuit for varying the potential impressed upon it from a range of negative values up through a series of positive values. This variation is secured by moving the point X along the potentiometer from point a to point b. At point a the grid may be made so strongly negative that no elec-

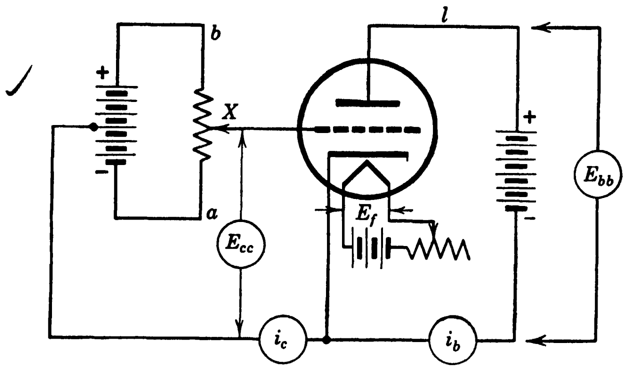


FIG. 9. Circuit for determining a triode transfer characteristic.

trons can pass to the plate. As the grid is made less negative, some electrons do pass to the plate, and with the movement of X to b the plate current i_b rises along the curve as shown in Fig. 10. After the grid becomes positive with respect to the cathode, a small current i_c begins to pass to the grid and follows the trend shown by the dotted curve on Fig. 10. It should be understood that the magnitude of the grid current is generally smaller than that indicated. The curve of Fig. 10 shows the operation of a tube for only one value of plate voltage, and, since a range of positive potentials may be used on the plate, the complete picture of the characteristic of a tube must be obtained through a family of curves for different plate potentials, as shown in Fig. 11 (lower right). A second set of family curves given in the upper right view of the same figure shows how the plate current varies with changes in the plate voltage. These curves are known as the plate characteristic curves, and they are very useful in designing electronic circuits. The plate characteristic curves may

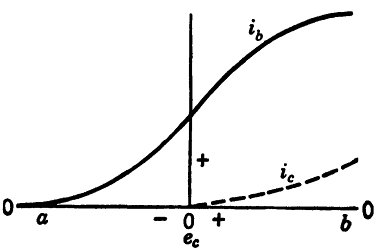
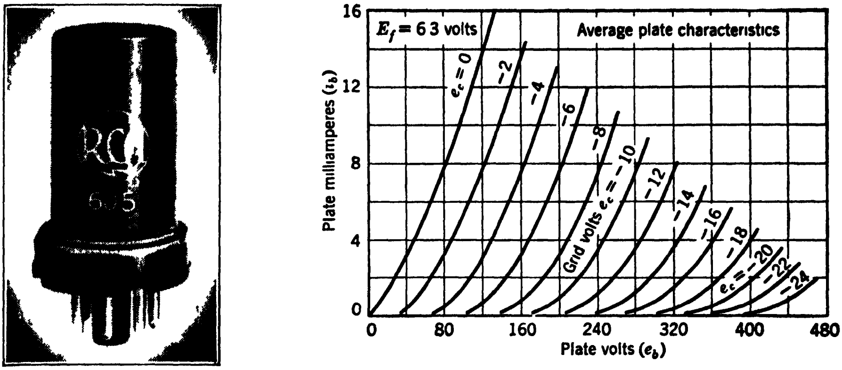


FIG. 10. Transfer characteristic of a triode.



DETECTOR-AMPLIFIER TRIODE

Heater—coated unipotential cathode	
Voltage	6.3 volts
Current	0.3 amp
Direct interelectrode capacitance	
Grid to plate	3.4 $\mu\mu$f
Grid to cathode	3.4 $\mu\mu$f
Plate to cathode	3.6 $\mu\mu$f
Maximum overall length	$2\frac{5}{8}$ in.
Maximum seated height	$2\frac{1}{16}$ in.
Maximum diameter	$1\frac{5}{16}$ in.

Amplifier

Plate voltage, max	300 volts
Grid voltage, min	0 volts
Plate dissipation, max	2.5 watts
D-c heater-cathode potential, max	90 volts
Cathode current, max	20 ma

Typical Operation—Class A_1 Amplifier

Plate	90	250 volts
Grid	0	−8 volts
Amplification factor	20	20
Plate resistance	6700	7700 ohms
Transconductance	3000	2600 μmhos
Plate current	10	9 ma

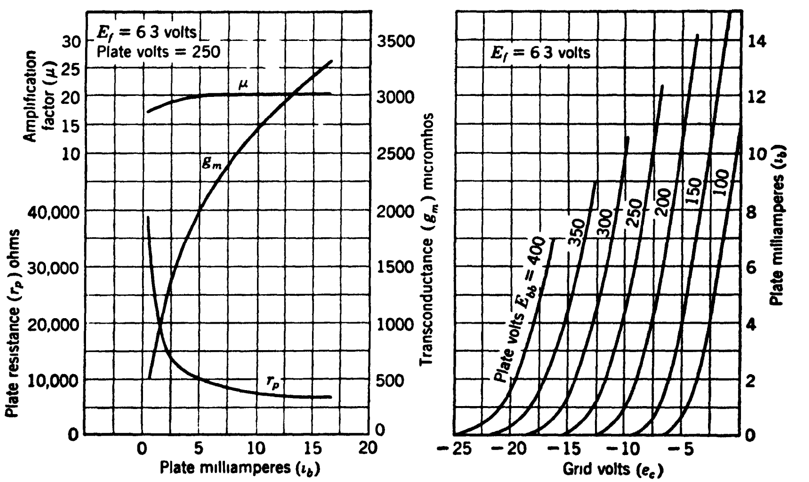


FIG. 11. Rating and characteristics of a detector-amplifier triode. (Courtesy Radio Corporation of America.)

be obtained directly by using a suitable test circuit, or they can be obtained by replotting the data given in the transfer characteristics.

Both sets of curves covered in the preceding paragraph are static characteristics, i.e, they apply for a static or constant potential between the cathode and the plate. In the application of the three-electrode

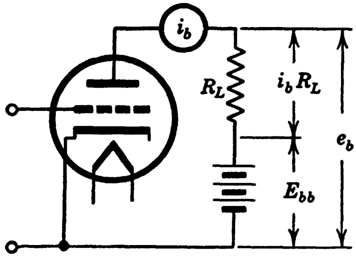


FIG. 12. Load circuit on a triode.

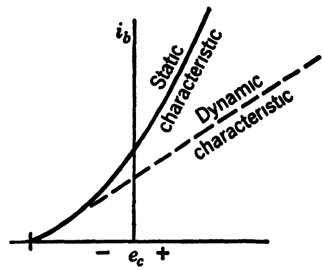


FIG 13 Static and dynamic characteristic of a triode.

tube some form of load must be placed in the plate circuit. This load (Fig. 12) will present a resistance R_L to the current and will give a fall of potential over itself. Thus the potential between the cathode and plate does not remain constant as E_{bb} but becomes e_b where $e_b = E_{bb} - i_bR_L$. This fall in plate potential will reduce the plate current

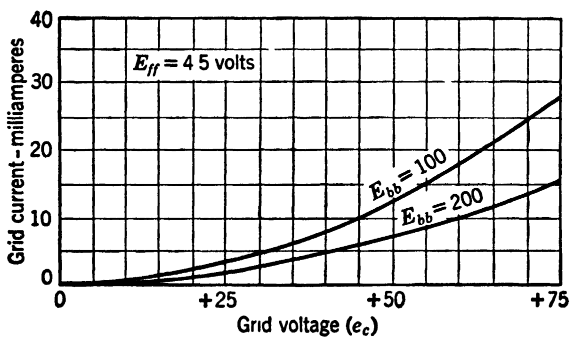


FIG. 14. Typical grid-current characteristics of a triode.

below the values determined by the circuit of Fig. 9. The effect of this lowering of plate potential with the increase of grid potential is to give a lower curve as shown dotted in Fig. 13. This new curve is the dynamic characteristic for the given load resistance. Obviously, the dynamic characteristics for a known load resistance can be determined

experimentally by inserting this load at point l on Fig. 9. However, since there are an infinite number of possible load resistances, it is not feasible to give all dynamic characteristic curves. Hence it is customary to determine the dynamic curve for a given load from a family of static curves by suitable calculations.

Two grid-current characteristics are given in Fig. 14. With a higher plate voltage E_{bb}, the plate attraction is relatively stronger and fewer electrons land on the grid, resulting in smaller values of grid current.

Amplification Factor. The chief value of the vacuum type of triode lies in its ability to amplify a relatively weak signal in the form of a change of potential or charge impressed across its grid. Such a change of potential will produce a rather large change of current in the plate circuit, and this change of current passing through a resistance or one winding of a transformer will produce a change of voltage of increased magnitude. It is possible for a mere change of charge on the grid (representing zero or nearly zero power input) to produce a large change of current and voltage in the plate circuit. Thus there would appear to be an infinite increase of power. In actual circuits some power is absorbed in the input circuits so that infinite amplification of power is not attained.

The amplifying power of a triode is the measure of the greater effectiveness of changes in the grid potential over those of the plate potential. The obvious way of stating such a measure is by means of a ratio of the voltages employed. Thus, mathematically, the amplification factor mu (μ) may be expressed as

$$\mu = -\frac{e_b - e_b'}{e_c - e_c'} \quad \text{(for } i_b = \text{constant)} \tag{1}$$

where $e_b - e_b'$ is the change in plate voltage required to compensate for a small change in grid voltage represented by $e_c - e_c'$. Since a decrease in plate voltage is necessary to compensate for an increase in grid voltage, the minus sign is necessary if mu is to be considered as a positive number. The factor mu may be determined from the transfer characteristics of the tube by substituting values in equation 1. Referring to Fig. 11 (lower right), it will be found that 11 milliamperes of plate current will be produced by 100 volts on the plate and a 0.0 volt on the grid. This same current will be produced by 200 volts on the plate and -5 volts on the grid. Thus

$$\mu = -\frac{(100 - 200)}{0 - (-5)} = \frac{100}{5} = 20$$

Similar calculations for other values of constant current will give a like value for mu. If points are taken for low values of plate current where the curvature of the characteristic is high, the value calculated for mu will vary.

In terms of calculus, the amplification factor is expressed as follows:

$$\mu = -\frac{\partial e_b}{\partial e_c} = -\frac{de_b}{de_c} \quad (i_b = \text{constant}) \tag{2}$$

The amplification factor depends on the geometry of the tube. The closer the control grid can be placed to the cathode, the more effective

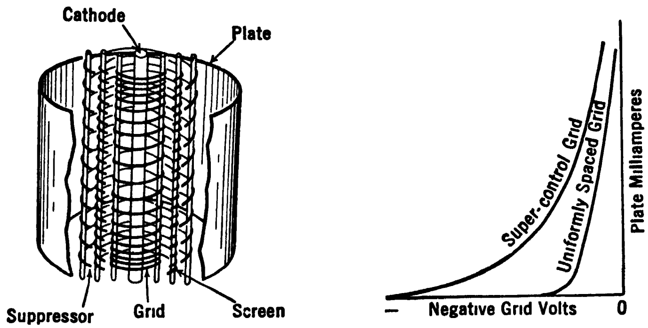


FIG. 15. Construction and curves of a variable-mu tube. (Courtesy Radio Corporation of America)

it will be in controlling the electron stream to the plate, and hence the higher the amplification factor will be. This closeness is limited by the necessary insulation of parts as well as by mechanical strength and construction. The amplification factor of a triode varies from about 3 as a minimum to about 100 as a practical maximum, the exact value depending upon the purpose for which the tube was designed.

For some applications it is desired to have a variable mu. This is secured by a close spacing of the grid wires at the top and bottom of the grid and a coarse spacing at the center. This construction and the effect upon the plate-current characteristic are illustrated in Fig. 15. The characteristic curve of a triode and other multielectrode tubes near the cutoff point is broadened, and the tube can be used for amplification at lower levels of plate current. This type of construction is known by the terms supercontrol grid, remote cutoff, and variable mu. Tubes employing this type of grid are used for automatic volume control.

Mutual Conductance and Transconductance. The mutual conductance of a triode is the rate at which the plate current changes with the grid voltage. Mathematically, it is the derivative of the plate current with respect to the grid voltage.

$$g_m = \frac{\partial i_b}{\partial e_c} = \frac{di_b}{de_c} \quad (e_b = \text{constant}) \tag{3}$$

Mutual conductance (g_m) is the slope of the transfer characteristic curve and is expressed in terms of microamperes per volt, or micromhos. The concept of mutual conductance may not be easy to grasp. It may help to think of the conductance of (one circuit) the plate-cathode circuit as being affected (mutually) by the voltage impressed across (a second circuit) the grid-to-cathode circuit.

For an example, take the point on the lower right-hand curves of Fig. 11 corresponding to 5 milliamperes on the 100-volt plate curve. Assume a current swing from 4.0 to 6.0 milliamperes which corresponds to a change of 1.0 volt on the grid. Hence

$$g_m = \frac{2 \times 10^3}{1} = 2000 \text{ micromhos}$$

Since the adoption of multigrids or electrodes for tubes, the term transconductance often displaces the expression mutual conductance. The mutual conductance referred to above becomes the control-grid-to-plate transconductance, and similar terms designate the electrodes involved in the transconductance in multielectrode tubes. The value of the transconductance varies because it depends on the varying curvature of the transfer characteristic of a tube.

Plate Resistance. The plate resistance of a triode is the rate of change of plate voltage with respect to the rate of change of plate current. Mathematically, it is the derivative of the plate voltage with respect to the plate current.

$$r_p = \frac{\partial e_b}{\partial i_b} = \frac{de_b}{di_b} \quad (e_c = \text{constant}) \tag{4}$$

Plate resistance is the reciprocal of the slope of the plate characteristic curves (Fig. 11, upper right). This plate resistance is frequently referred to as the a-c resistance or the dynamic resistance to distinguish it from the d-c plate resistance. The d-c plate resistance is the quotient of the static plate to cathode voltage E_{bo} divided by the static or quiescent plate current I_{bo}. The d-c plate resistance is of little im-

portance in tube operation. For a sample calculation of a-c plate resistance r_p, use the curve $e_c = -4.0$ and $i_b = 10$ milliamperes for the upper right curves of Fig. 11. Here a swing of 40 volts on the plate will give a current change of 5.3 milliamperes. Thus

$$r_p = \frac{40}{5.3 \times 10^{-3}} = 7500 \text{ ohms} \quad \text{(approx)}$$

Parameters of Multielectrode Vacuum Tubes. The three factors of amplification—constant μ, control-grid-to-plate transconductance g_m, and plate resistance r_p—are termed the parameters of multielectrode vacuum tubes. These factors are related to each other as follows:

$$r_p \times g_m = \mu \tag{5}$$

This relationship is proved by a substitution of the expression for each of these three factors from equations 2, 3, and 4.

$$\frac{de_b}{di_b} \times \frac{di_b}{de_c} = \frac{de_b}{de_c}$$

The magnitude of the parameters depends on the geometry of the tube, such as the spacing of the cathode, grid, and plate, the diameter of the grid wires, the spacing of the grid wires, and the area of the plate. The parameters of a tube are important in the design of vacuum-tube circuits.

The parameters of a triode vary with plate current as illustrated in Fig. 11 (lower left) for a 6J5 triode. The amplification factor is usually constant throughout a wide range of plate current, but the transconductance and plate resistance usually show much variation. In the design of circuits the values of the parameters are obtained from data furnished in the tube manuals of the manufacturers, or they may be calculated from test data using the methods suggested in the preceding discussion.

Plate Current in a Triode. Childs' equation for the two-electrode tube can be modified slightly for application to the triode. The modified equation becomes

$$i_b = K\left(\frac{e_b}{\mu} + e_c\right)^{3/2} \tag{6}$$

where K is a constant depending on the tube dimensions. Obviously, the expression in the parenthesis refers to the grid potential and not to the plate potential of Childs' equation. This equation does not hold for low values where mu is not constant nor for high values where

saturation effects are present; also, for positive values of grid potential, i_b represents the sum of plate and grid currents.

Types of Triodes. Triodes may be classified on the basis of (1) their construction, (2) their use, or (3) their power rating. Under the first classification triodes are built with all glass enclosures, in metal envelopes with air cooling, and in metal tubes with water cooling. Most tubes use glass enclosures. Glass construction gives a lower cost because it utilizes the manufacturing technique developed through years in producing electric light bulbs. Glass has a low heat-dissipating ability and is limited to use in tubes having a maximum plate dissipation of approximately 1000 watts. Two kinds of glass are used in forming glass tubes—an ordinary soft glass for small tubes and for tubes of low rated capacity, and a hard glass having a higher softening and melting temperature for power tubes, which must dissipate considerable heat and operate at a high temperature.

Two glass triodes which have wide use as amplifiers for repeaters on telephone toll lines are shown in Fig. 16. A low-power triode having a glass enclosure and a graphite plate is illustrated in Fig. 17. Its moderately high grid- and plate-voltage rating makes it adaptable for use in transmitter circuits as an amplifier, oscillator, or modulator. Interesting glass triodes are shown in Fig. 18. These tubes are very small and compact and are designed for ultra-high-frequency service. For this service it is necessary to reduce the interelectrode capacities to a minimum, which means small electrodes and short lead-in conductors.

The metal receiving tube is small in physical size and its metal case serves as a shield from external fields. The active parts of the tube are of normal size, but a large saving in space is made in the stem (see Fig. 19) and in the metal envelope which fits the electrode assembly closely. The metal tube utilizes a new alloy called Fernico and a glass which have practically the same coefficient of expansion throughout a wide range of temperature. The Fernico base of the metal tube consists of a flanged disk in which a number of eyelet holes are pressed. Glass beads are welded into these eyelets so as to support the leads and insulate them from each other and the metal case (Fig. 19). A small metal tube is welded into the center of the disk and serves as the medium for exhausting the assembled electron tube. After the electrodes are assembled on the base, the metal cover is welded to the base in a single quick butt weld. Then the tube is exhausted and a pinch weld on the exhaust tube seals the triode.

The amount of heat energy that can be dissipated by radiation and convection in air from glass and metal tubes is rather limited because

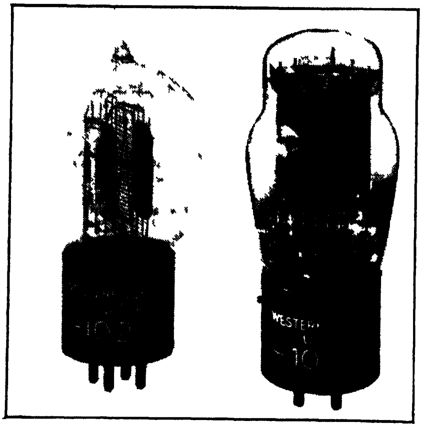


Fig. 16. Two filamentary triodes for telephone service *left*, voltage-amplifier; *right*, low-power amplifier for voice and carrier frequency. (Courtesy Western Electric Company.)

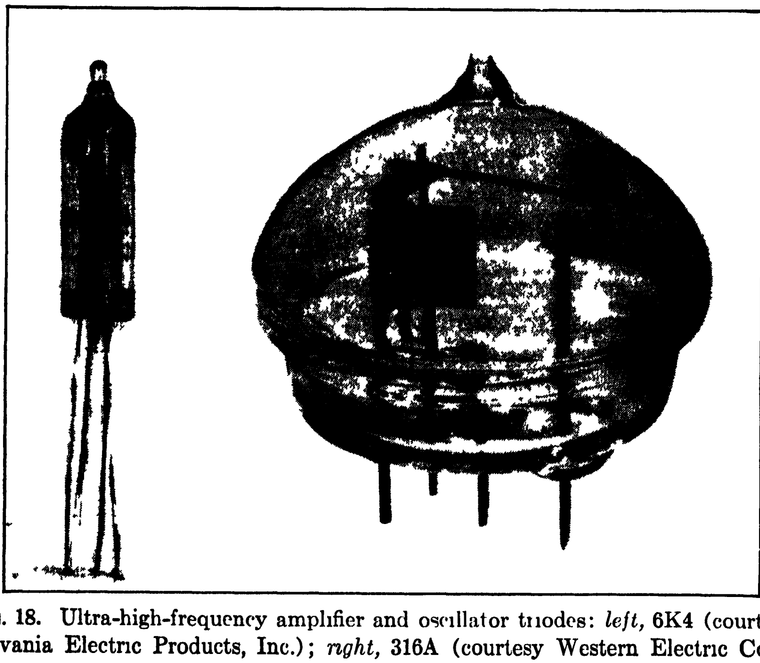
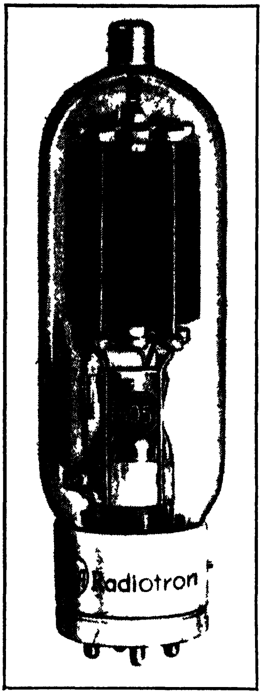


Fig. 18. Ultra-high-frequency amplifier and oscillator triodes: *left*, 6K4 (courtesy Sylvania Electric Products, Inc.); *right*, 316A (courtesy Western Electric Company).



R-F POWER AMPLIFIER AND OSCILLATOR

General Data

Filament—thoriated tungsten	
Voltage (a-c or d-c)	10 volts
Current	3.25 amp
Direct interelectrode capacitances (approx)	
Grid to plate	6 5 μμf
Grid to filament	8.5 μμf
Plate to filament	10.5 μμf

Class C Telegraphy

D-c plate voltage, max	1500 volts
D-c grid voltage, max	−500 volts
D-c plate current, max	210 ma
D-c grid current, max	70 ma
Plate input, max	315 watts
Plate dissipation, max	125 watts

Typical Operation

Filament voltage	10	10	10 a-c volts
D-c plate voltage	1000	1250	1500 volts
D-c grid voltage	−95	−100	−105 volts
Peak r-f grid voltage	225	230	235 volts
D-c plate current	200	200	200 ma
D-c grid current, approx	40	40	40 ma
Driving power, approx	8 5	8 5	8 5 watts
Power output, approx	130	170	215 watts

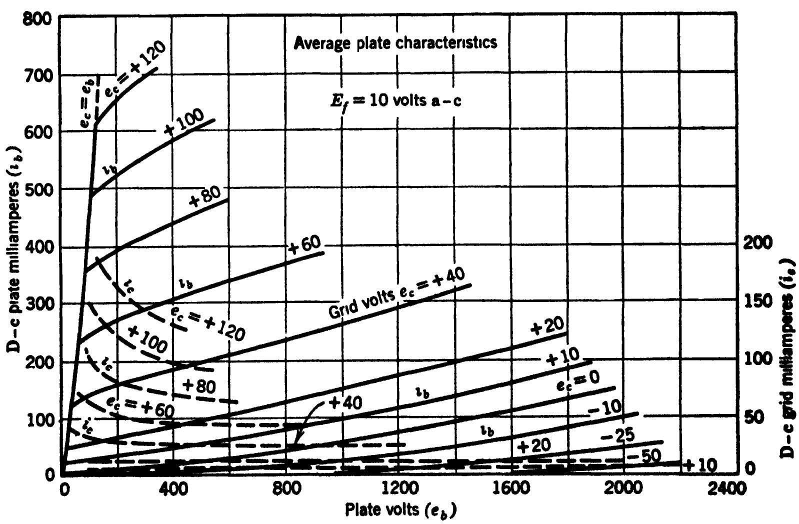


Fig. 17. Rating and characteristics of a low-power vacuum triode. (Courtesy Radio Corporation of America.)

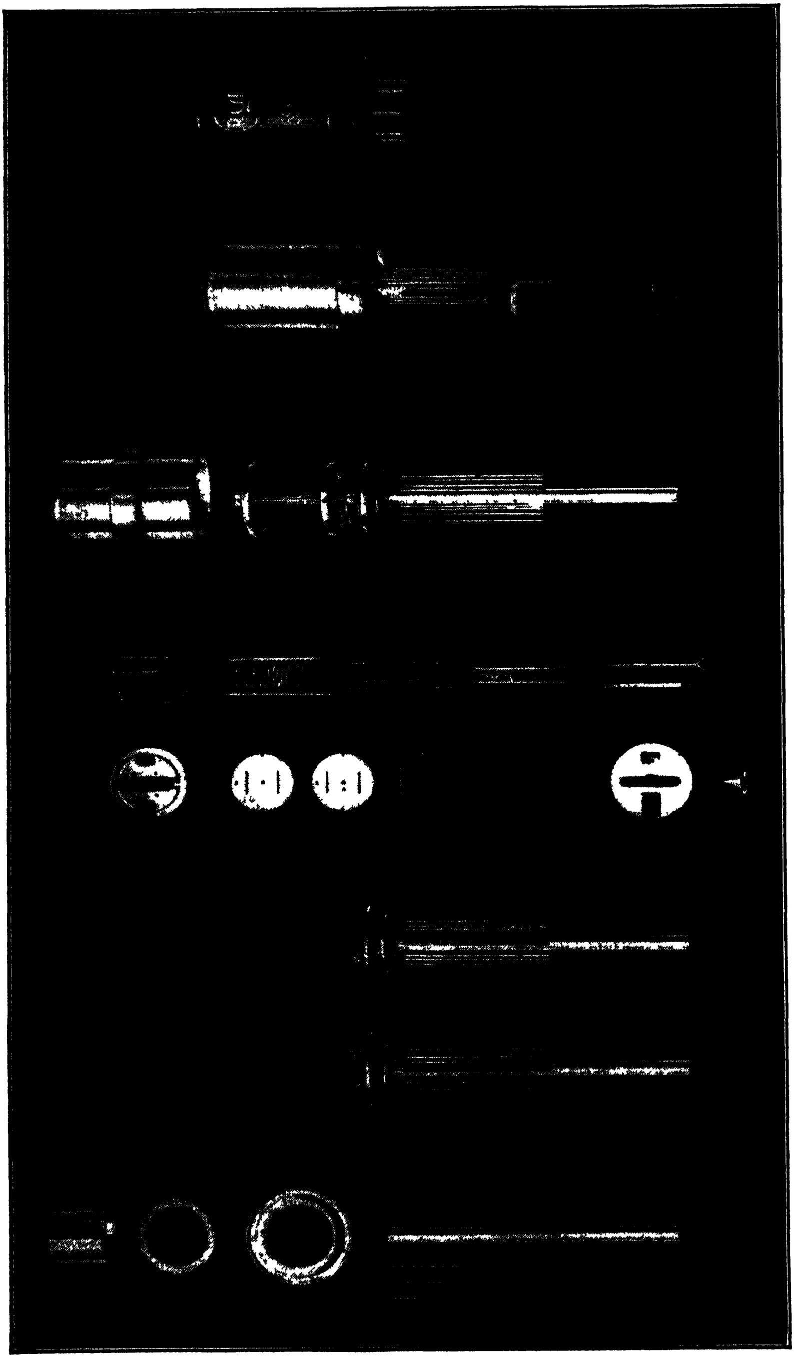


FIG 19 Parts and assembly of a metal receiving tube. (Courtesy Radio Corporation of America)

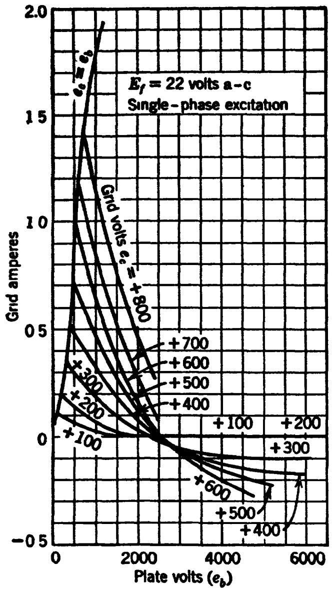
the volume of the tube increases more rapidly than the surface area. It is possible to obtain a greater output by operating glass tubes in parallel. A more direct solution is the use of water cooling or forced air cooling for tubes. In the water-cooled triode the plate is a round copper tube, which also constitutes the envelope of the tube as shown in Fig. 20. The copper plate is immersed in a bath of circulating water which conducts the heat away rapidly. It is sealed at the top to a glass envelope which serves to insulate and support the cathode and grid of the tube. This airtight, metal-to-glass union is called the housekeeper seal. In making this seal the open end of the copper plate is drawn out to a feather edge and placed in the molten glass of the envelope. Although there is some difference in the coefficient of expansion of the glass and copper, the seal along the thin part of the feather edge is seldom broken.

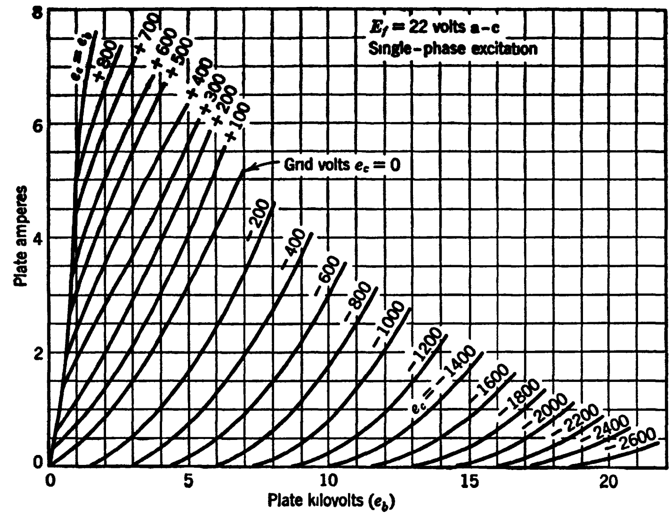
Water-cooled tubes are usually built in capacity ratings of from 1 to 100 kilowatts and plate-voltage ratings from 5000 to 100,000 volts. Tubes of 250-kilowatt rating have been built in the United States for use abroad. The rating and characteristics of a 10-kilowatt water-cooled triode are given in Fig. 20.

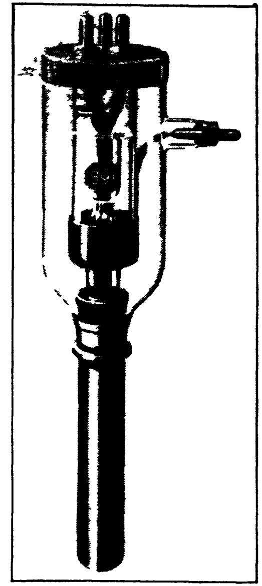
Water cooling of tubes involves some construction and maintenance problems, which has led to the development of power tubes using forced air cooling. A large capacity and moderately high-voltage tube of this type is illustrated in Fig. 21. The student should compare the construction details, rating, and characteristics of this tube with the preceding water-cooled unit.

The three-electrode vacuum tube may be classified as amplifier and relay, oscillator or alternating-current generator, modulator, and detector or demodulator. The particular function performed by the tube depends not upon its construction but upon the external circuits with which it is associated. Thus a single tube might be connected into circuits for serving as amplifier, oscillator, modulator, or detector. The theory of amplification is illustrated in Fig. 22 (right). In practice certain types of tubes are chosen for the different applications, but this selection is not determined by any difference in their inherent theory of operation. The circuits and applications of triodes will be covered in succeeding chapters.

Limitations of a Triode. There are two properties of a triode and its circuit that limit its range of operation as an amplifier. The first of these arises from the capacitance coupling between the grid and the plate. The grid and the plate are two electrodes separated by an







R-F POWER AMPLIFIER AND OSCILLATOR

General Data

Filament—two-section tungsten	
Voltage per section	11.0 volts
Current	60.0 amp
Amplification factor	8.0
Interelectrode capacitance	
Grid to plate	27.0 μμf
Grid to filament	18.0 μμf
Plate to filament	2.0 μμf
Mounting position—filament end up	

Class C Telegraphy

Maximum Ratings, Absolute Values

D-c plate voltage, max	12,000 volts
D-c grid voltage, max	−3,000 volts
D-c plate current, max	2.0 amp
D-c grid current, max	0.15 amp
Plate input, max	18 kw
Plate dissipation, max	6 kw

Typical Operation

D-c plate voltage	8,000	10,000 volts
D-c grid voltage	−1,800	−2,000 volts
Peak r-f grid voltage	2,500	2,900 volts
D-c plate current	1.1	1.45 amp
D-c grid current, approx	0.06	0.11 amp
Driving power, approx	150	310 watts
Power output, approx	6.5	10 kw

FIG 20. Rating and characteristics of a water-cooled transmitting triode. (Courtesy Radio Corporation of America.)

TRANSMITTING TRIODE

Forced Air Cooled

Electrical

Filament—three-section tungsten	
Excitation—1ϕ a-c, 3ϕ a-c, 6ϕ a-c, or d-c	
Voltage per strand	10 volts
Current per terminal	61 amp
Amplification factor	36
Direct interelectrode capacitances (approx)	
Grid to plate	34 $\mu\mu$f
Grid to filament	48 $\mu\mu$f
Plate to filament	3 5 $\mu\mu$f

Maximum Ratings, Absolute Values

D-c plate voltage, max	20,000 volts
D-c grid voltage, max	−3,000 volts
D-c plate current, max	4 amp
D-c grid current, max	0 4 amp
Plate input, max	70 kw
Plate dissipation, max	20 kw
Radiator temperature, max	180° C

Typical Operation

D-c plate voltage	12 000	15 000	18,000 volts
D-c grid voltage	−800	−900	−1,000 volts
Peak r-f grid voltage	1 430	1 520	1,630 volts
D-c plate current	3 5	3 6	3 6 amp
D-c grid current, approx	0 26	0 25	0 21 amp
Driving power, approx	360	370	340 watts
Power output, approx	30	40	50 kw

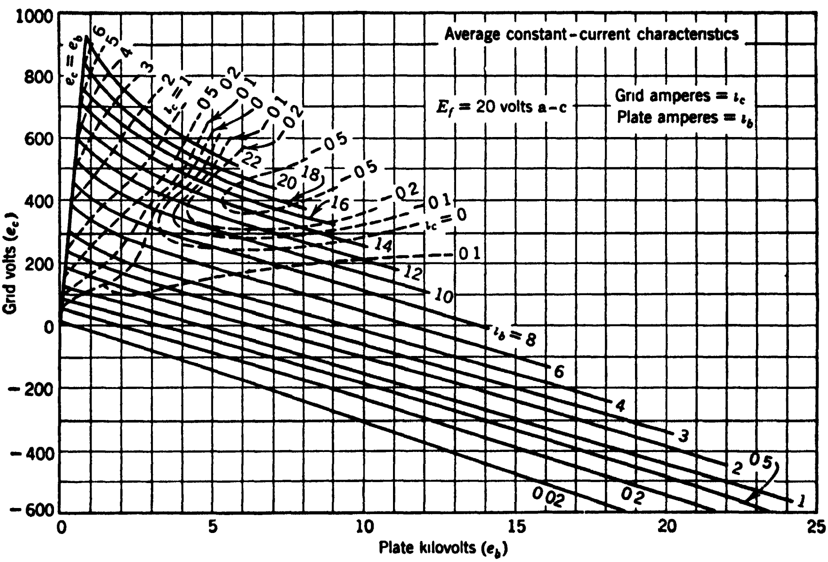


Fig 21 Rating and characteristics of a power triode (Courtesy Radio Corporation of America)

insulator (a vacuum), and thus they constitute a condenser. A condenser is an open circuit to a direct current but a conductor for alternating currents. Thus, although the cathode-grid and cathode-plate circuits of Fig. 22 (left) appear to be separated, the interelectrode capacity between plate and grid forms a coupling between them. Through this coupling a feedback may occur from plate to grid caused by changes of potential (alternating) on the plate. This feedback will

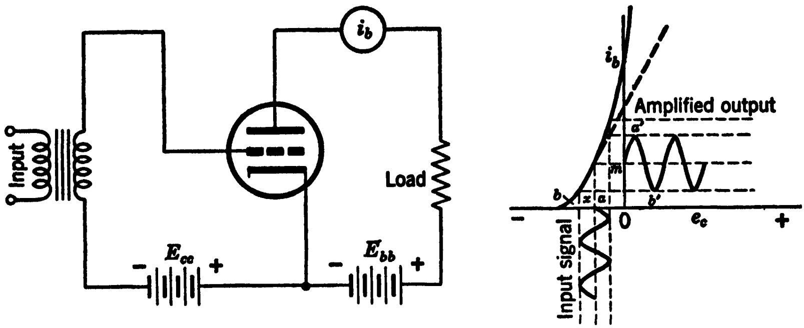


FIG. 22. Simple triode circuit and diagram illustrating the theory of amplification.

interfere with the normal control by the grid, or it may even start oscillations and cause the tube to serve as an oscillator.

The second limitation in the triode amplifier circuit results from the changes in the cathode-plate potential with the changes in the plate current. In order to utilize the amplified signal, it is necessary to place a resistance, inductance, or transformer as a load in the plate circuit. (See Fig. 22.) This load has impedance and the plate current through this impedance produces a voltage drop ZI. A varying plate current produces a varying drop over the load impedance, which in turn is subtracted from the constant potential supply of the plate E_{bb}. Obviously, changes in cathode-plate potential will vary the plate current as was pointed out in the discussion of the two-electrode tube (see Fig. 9, Chapter III, page 47). This change in plate current is a direct action on the plate-cathode circuit, whereas the interelectrode action suggested in the preceding paragraph takes place in the grid-cathode circuit.

The Tetrode. The limitations in the use of the triode as outlined in the preceding article can be overcome by the addition of a fourth electrode, a grid, placed between the regular or control grid and the plate. This fourth electrode is called the screen grid, and the tube is known as a screen-grid tube.

The construction of the screen-grid tube is shown schematically in Figs. 23 and 24. In the smaller receiving tube the screen grid consists of two grids connected at the top to form a nearly complete screen or shield for the plate. This construction reduces the interelectrode capacity between grid and plate to an average of $\frac{1}{800}$ of its value without the screen and overcomes the first limitation of the triode.

A typical circuit for using a screen-grid tube is given in Fig. 24. The screen grid is connected to the cathode and is maintained at a constant potential lower than the normal plate voltage. Thus the screen grid sets up in the space surrounding it a constant potential. This potential supplies a constant accelerating force (attraction) for any electrons that pass the control grid. Normally nearly all electrons that reach the position of the screen grid pass through the meshes of the grid and land on the plate. *Moderate changes in the plate potential will not influence the number of electrons reaching the plate.* Thus the screen has served to overcome the second limitation of the triode. The screen grid attracts and gathers those electrons which are traveling directly toward its mesh wires, and these electrons constitute a small screen-grid current.

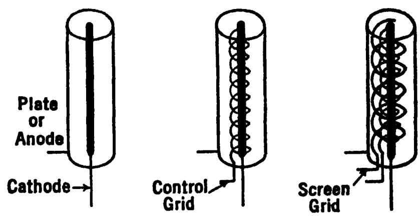


FIG. 23. Location of the screen grid in a vacuum tube.

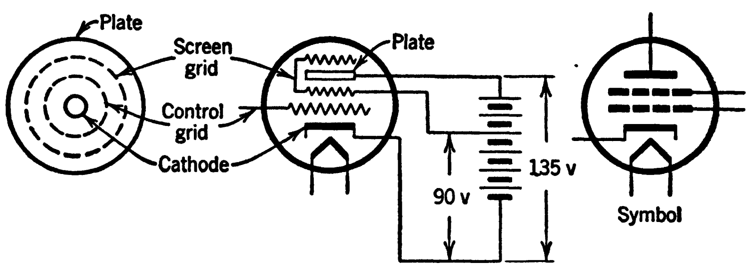


FIG. 24. Schematic diagram of a tetrode and its circuit.

The action of the screen grid may be pictured from the potential distribution curve of Fig. 25. The cathode is held constant at zero (ground potential) while the space surrounding the screen grid is held at another constant potential (say 90 volts). The control-grid potential is governed by the incoming signal and may swing from negative to positive. Likewise, the plate potential may swing from 90 to 135

volts. These latter swings will not disturb the constant potential at the screen grid and will not affect the electron flow to the plate. It should be noted that the actual potential distribution in a tetrode is not so simple as that indicated in Fig. 25. Space charges near the cathode or other electrodes, the geometry of the tube, and the intergrid wire regions will change the linear distribution which has been used to give a simple concept of the operation of the tube.

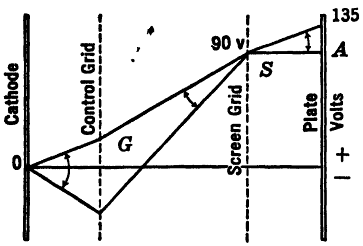


FIG. 25. Potential distribution in a screen-grid tube.

The screen-grid tube gives an unexpected performance when the plate potential falls to low values. This performance can be determined by the circuit of Fig. 26 (right), from which the plate current can be determined for a variation of plate potential from zero to maximum rated value. The early types of screen-grid tubes (Type 24-A) gave the plate-current plate-voltage characteristic shown in Fig. 26 (left), when the screen-grid voltage was held at normal value and the control grid kept at zero volts (cathode potential). At zero plate potential, the plate current is zero, as would be expected. Then, as plate potential rises to 10 or 15 volts, the plate current rises. With a further rise in plate potential, the current starts to drop and continues to do so until it becomes negative in value.

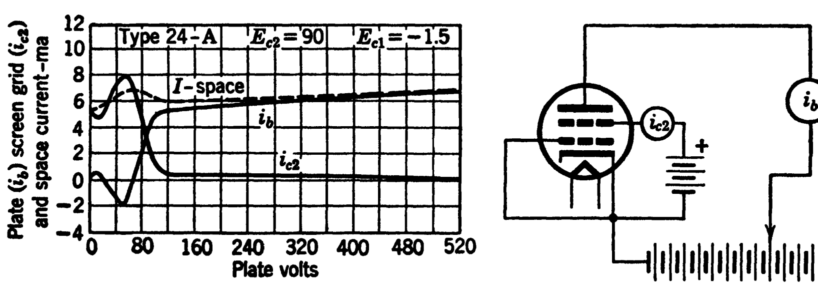


FIG. 26. Plate-current-voltage characteristic and circuit of a screen-grid tube.

After reaching a minimum point, the plate current rises with plate potential to a normal value and then becomes approximately constant for further rise in plate voltage. This unusual characteristic is due to secondary emission of electrons from the plate which are attracted to the screen grid. Secondary emission occurs regularly at the plate of

all tubes owing to impinging electrons, but for the triode all the electrons of secondary emission are attracted back to the plate and hence do not produce any change in the current of the plate circuit. In the screen-grid tube, electrons are splashed out of the plate as shown in Fig. 27. If the screen grid happens to be at a higher potential than the plate and if the electrons are splashed out with sufficient velocity, some or all of these electrons of secondary emission will pass to the screen grid. If there are more electrons of secondary emission than original impinging electrons, the electron current may reverse in the plate circuit and flow to the screen grid. Thus the characteristic of Fig. 26 becomes easy to understand. Beginning at the left (zero plate potential), no electrons are attracted by plate. Then, as plate potential goes somewhat positive, electrons are attracted to it but without sufficient velocity to cause secondary emission. With a rise in plate potential a point is soon reached where secondary emission occurs with some electrons going to the screen grid. A further increase in plate potential increases the secondary emission rapidly and with it the number of electrons going to the screen grid. As the potential on the plate approaches that of the screen grid, the plate begins to receive back more electrons and a reversal of plate current trend takes place. When the plate potential passes the screen in magnitude, the plate current rises to full or normal value.

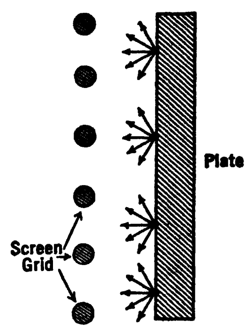
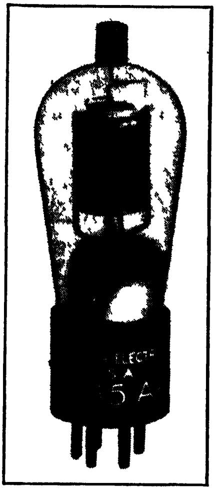


FIG. 27. Secondary emission in a tetrode.

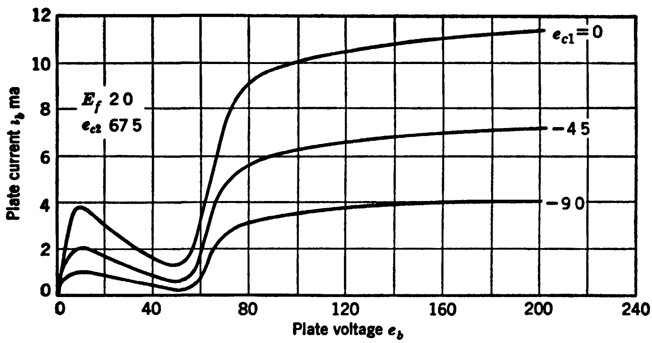
The screen-grid current for the tube in question is shown by the i_{c2} curve of Fig. 26. This curve is the inverse of the plate current, as might be expected. The space current or sum of plate and screen-grid current is approximately constant as shown by the curve of Fig. 26, because the electrostatic field produced by the screen grid is constant.

Screen-grid tubes of later manufacture have been improved by special treatment of the plate which reduces their tendency toward secondary emission. This improvement prevents the plate current from becoming negative and gives characteristic curves as shown in Fig. 28. It is evident from these characteristic curves that the screen-grid tube should give more linear amplification as long as the plate potential exceeds the screen-grid potential, and that wide swings in plate potential may occur in that region. For lower values of plate potential the tube will give unsatisfactory service. The effect of secondary



SCREEN-GRID VOLTAGE AMPLIFIER

Filament	
Voltage	2 0 volts
Current	1 6 amp
Amplification factor	85–170
Interelectrode capacitance	
Grid to plate	0 025 $\mu\mu$f
Input	4 5 $\mu\mu$f
Output	8 0 $\mu\mu$f
D-c plate voltage	135–180 volts
D-c screen voltage	45–67 5 volts
D-c grid bias	1 5–4 5 volts



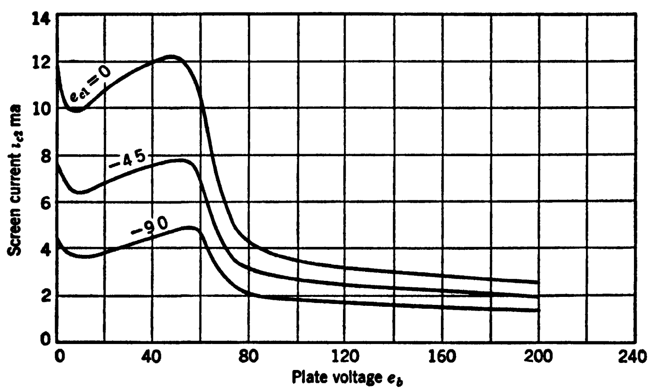


FIG. 28 Rating and characteristics of a voltage amplifier screen-grid tetrode. (Courtesy Western Electric Company)

emission in the screen-grid tube can be overcome in two different ways, to be covered in succeeding articles. The principal application of screen-grid tubes is for radio-frequency amplifiers. The rating and characteristics of a typical screen-grid tube are given in Fig. 28.

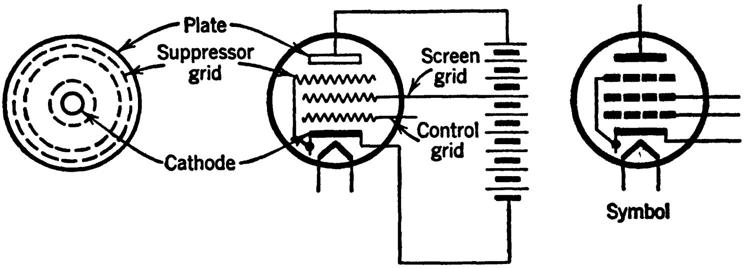


FIG. 29. Schematic diagram and circuit of a pentode.

The Pentode. The pentode is a tube having five electrodes. The fifth electrode, called the suppressor grid, was added to overcome the harmful effect of the secondary emission in the screen-grid tube. The suppressor grid is a mesh placed between the plate and screen grid as shown in Fig. 29. It is usually connected directly to the cathode inside the tube, though it may be brought outside for other connections. When connected to the cathode the supressor grid serves as a shield or suppressor to prevent the electrons of secondary emission from passing to the screen grid. The suppressor grid performs this function by creating a field of near zero potential through which electrons (secondary emission) must pass before coming under the influence of screen grid. This explanation may become more clear from a study of Fig. 30, which gives a schematic potential distribution for a pentode. Here point P represents the constant zero potential on the suppressor grid. Electrons splashed out of the plate must move away from it against the attraction of its positive potential and past point P before they come under the influence of the potential on the screen grid.

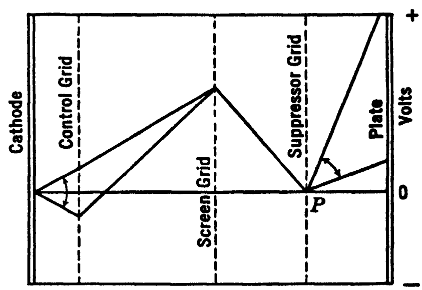
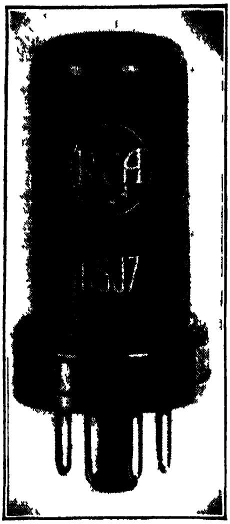


FIG. 30. Potential distribution in a pentode.



TRIPLE-GRID DETECTOR-AMPLIFIER

Heater—coated		
Voltage (a-c or d-c)		6.3 volts
Current		0.3 amp
Direct interelectrode capacitance		
Pentode Conn.	Grid to plate, max	0 005 μμf
	Input, max	6.0 μμf
	Output, max	7.0 μμf
Triode Conn.	Grid to plate, max	2.8 μμf
	Grid to cathode, max	3.4 μμf
	Plate to cathode, max	11 μμf

AMPLIFIER (PENTODE CONNECTION)

Plate voltage, max	300 volts
Screen voltage, max	125 volts
Screen supply voltage, max	300 volts
Grid voltage, min	0 volts
Plate dissipation, max	2.5 watts
Screen dissipation, max	0.3 watt

Typical Operation and Characteristics—Class A_1 Amplifier

Plate	100	250 volts
Screen	100	100 volts
Grid	−3	−3 volts
Suppressor—connected to cathode at socket		
Plate resistance	0.7	meg
Transconductance	1575	1650 μmhos
Grid bias (for plate current = 10μamp)	−8	−8 volts
Plate current	2.9	3 ma
Screen current	0.9	0.8 ma

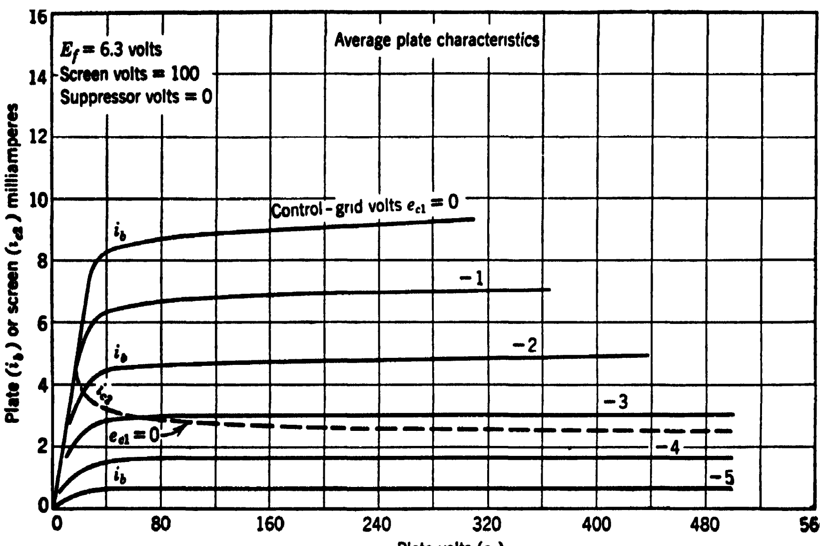


FIG. 31. Rating and characteristics of a pentode detector-amplifier. (Courtesy Radio Corporation of America.)

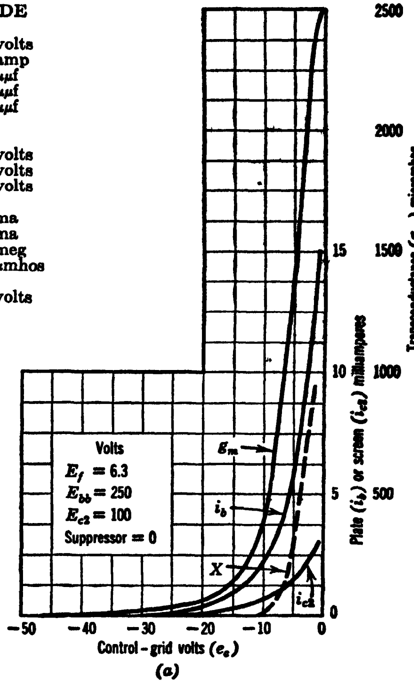
TRIPLE-GRID SUPERCONTROL PENTODE

Heater voltage (a-c or d-c)	6.3 volts
Heater current	0.3 amp
Grid-plate capacitance,* max	0.003 $\mu\mu$f
Input capacitance *	6 $\mu\mu$f
Output capacitance *	7 $\mu\mu$f

Typical Operation

Plate voltage	100	250 volts
Screen voltage	100	100 volts
Grid voltage	−3	−3 volts
Suppressor—connected to cathode at socket		
Plate current	8.9	9.2 ma
Screen current	2.6	2.4 ma
Plate resistance, approx	0.25	0.8 meg
Transconductance	1900	2000 μmhos
Grid bias for transconductance of 10 μmhos	−35	−35 volts

* With shell connected to cathode.



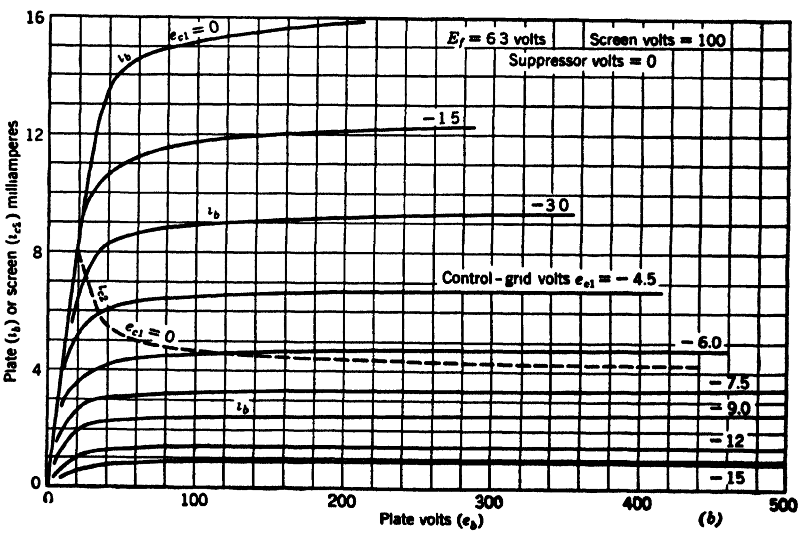


FIG. 32. Rating and characteristics of a triple-grid super-control pentode. (Courtesy Radio Corporation of America.)

The plate-current plate-potential characteristics of pentodes are given in Figs. 31 and 32. The pentode permits wide swings of plate potential without affecting the fidelity of amplification. The action of the suppressor permits large power output for low input voltage on the control grid and it permits high-voltage amplification at moderate values of plate voltage. Accordingly, some pentodes have a large power output and are used in the final stage of audio amplification for

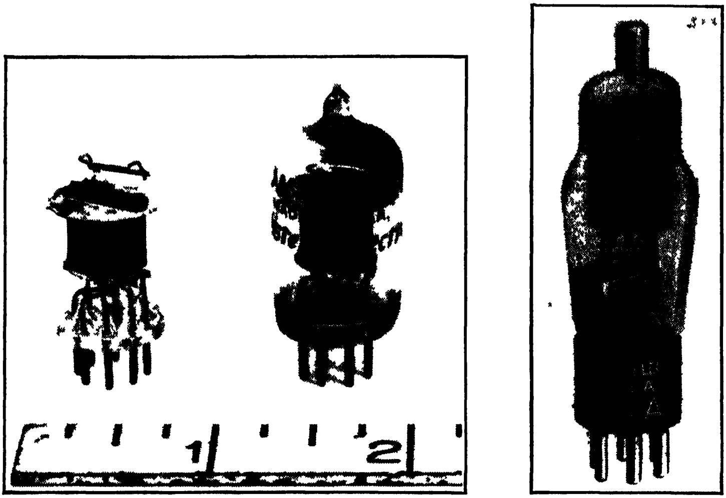


FIG. 33. Pentodes. Those on the left are for low-power applications at high and ultra-high frequencies (Courtesy Western Electric Company.)

supplying the current to the loudspeaker. Other types of pentodes are used as voltage amplifiers. The construction of typical pentodes is illustrated in Figs. 31, 32, and 33. The rating and characteristics of two pentodes are given in Figs. 31 and 32.

In both pentodes and tetrodes, the control grid is sometimes designed to give a variable mu. This is secured by a close spacing of the grid wires at the top and bottom of the grid and a coarse spacing at the center. The effect of this construction upon the plate-current characteristic is illustrated in Fig. 32 (upper right). The characteristic curves of the tube near the cutoff point are broadened and the tube can be used for amplification at lower levels of plate current. For comparison purposes a curve for a normal grid is shown by the dotted line marked X. Tubes employing the variable-mu type of grid are used for grid-bias volume control in radio-frequency amplifiers.

Beam Power Tube. A beam power tube is a tetrode * in which directed electron beams increase the power capacity and operating characteristics of the tube. This tube has three special features in its construction. First, the screen and control grid are composed of wires wound in helices so that each turn of the screen is shaded from the cathode by a turn of the control grid. This careful alignment tends to pass the electrons to the plate in beams and serves to reduce the magnitude of the screen-grid current. The second feature in the construc-

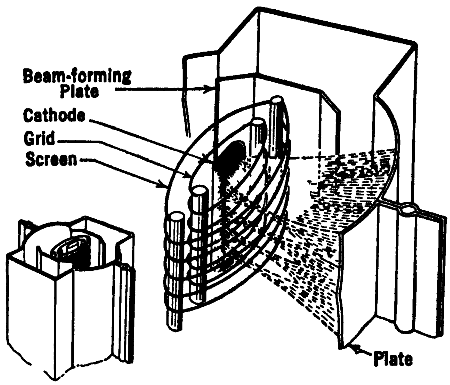


Fig. 34. Construction of a beam power tube.

tion of the beam power tube is the use of beam-forming plates which are connected to the cathode (Fig. 34). These plates serve to prevent any electrons from leaving the grid near its end supports and serve to give a sharp cutoff for the beams. The third feature of construction of the beam power tube is that the screen grid and plate are spaced relatively far apart (Fig. 34) and that these electrodes are operated at approximately the same potential. This construction results in a suppressor action between the screen and plate. To understand this action, assume that screen and plate are at the same potential and that the cathode is cold (zero emission). Under this condition the potential distribution between the screen and plate will be uniform and no potential gradient will exist. Next assume similar conditions with normal emission from the cathode and a control-grid potential that permits a normal plate current. Now moving electrons are present everywhere between screen and plate, and these electrons constitute a negative space charge which is strengthened by the action

* If beam-forming plates are called electrodes, the tube may be considered a pentode.

of the suppressor plates and by the fact that the electrons exist in concentrated beams. This negative space charge lowers the potential in the region between screen and plate and results in *a change of pace of the moving electrons.* This change of pace results in a variation of density of the electrons in the screen-plate space with a corresponding change of potential distribution. The net result of this action is the formation of a virtual suppressor between the screen and plate as shown in Fig. 35.

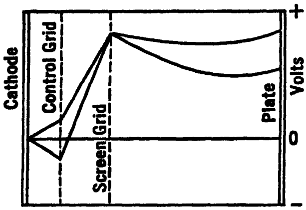


FIG. 35. Potential distribution in a beam power tube.

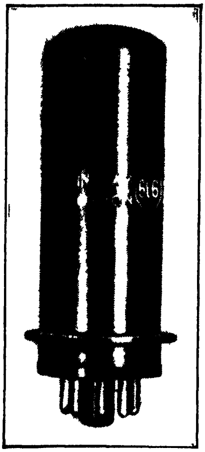
The rating and characteristics of a typical beam power tube are given in Fig. 36. It should be noted that wide changes of plate potential above 75 volts have little effect upon plate current. Also, the lower set of curves shows a sharp bend at the knee of plate-current plate-voltage curve. This means that little harmonic distortion occurs above the bend.

The advantages of tubes of the beam power type are high power output, high power sensitivity, and high efficiency. They are frequently used in the output stage of radio receivers and for other forms of load.

Pentagrid Converter. The pentagrid converter is a multielectrode tube having one cathode, one plate, and five grids. It is a dual-purpose tube which does not involve any new theory of action above that covered in preceding types. It is commonly used as a frequency converter, and the use of five grids gives rise to the name pentagrid converter. Frequency conversion involves the action of an oscillator and a modulator. The circuits for the pentagrid converter combine these two functions within a single tube, and the electron stream serves as the coupling.

Cathode-Ray Tube. A cathode-ray tube is a device for producing electron beams and for projecting them upon a fluorescent screen to give a picture of some electrical phenomenon. The first device of this type was developed by Braun in 1897. Cathode-ray tubes have used (1) cold electrodes with high potentials in vacuum, (2) hot cathodes with low gas pressure and low voltage between electrodes, and (3) hot cathodes in vacuum with fairly high accelerating potentials. Nearly all the tubes in use today fall in the third class.

A typical cathode-ray tube with electrostatic controls embodies the schematic construction shown in Fig. 37. The seven electrodes in this device serve to produce electrons, to concentrate them in a small beam, and to aim that beam upon various parts of the fluorescent screen on



BEAM POWER AMPLIFIER

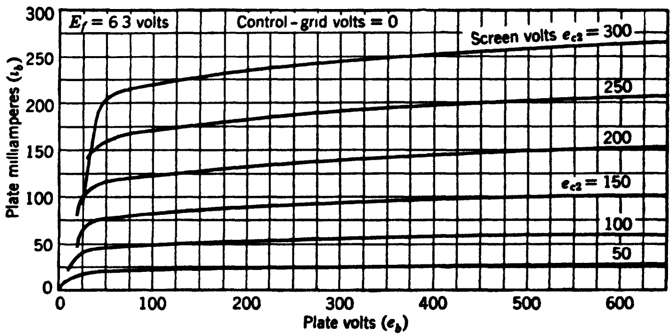
Heater—coated unipotential cathode	
Voltage (a-c or d-c)	6 3 volts
Current	0.9 amp

SINGLE-TUBE AMPLIFIER—CLASS A_1

Plate voltage, max	360 volts
Screen voltage, max	270 volts
Plate dissipation, max	19 watts
Screen dissipation, max	2 5 watts

Typical Operation

	Fixed	Bias	Cathode	Bias	
Plate	250	350	250	300	volts
Screen	250	250	250	200	volts
Grid	−14	−18			volts
Cathode resistor			170	220	ohms
Peak a-f grid voltage	14	18	14	12 5	volts
Zero signal plate current	72	54	75	51	ma
Max signal plate current	79	66	78	54 5	ma
Zero signal screen current	5	2 5	5 4	3	ma
Max signal screen current	7 3	7	7 2	4 6	ma
Plate resistance	22 500	33 000			ohms
Transconductance	6 000	5 200			µmhos
Load resistance	2,500	4,200	2 500	4,500	ohms
Total harmonic dist	10	15	10	11	%
Max signal power output	6 5	10 8	6 5	6 5	watts



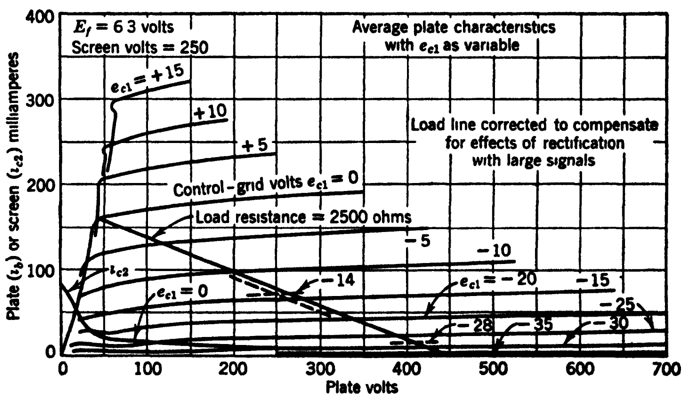


FIG. 36 Rating and characteristics of a beam power amplifier. (Courtesy Radio Corporation of America)

the end of the tube. The first five electrodes on the left constitute what is called the electron gun. The first electrode on the left is the cathode, a barium-coated plate with a heater behind it. Electrons from the cathode pass through a small hole in the control grid. Next, these electrons are accelerated by a higher potential on the second grid which is a disk with a small hole at its center. The first anode is a tube with two baffles, each containing a tiny hole at the center. This first anode has a potential of several hundred volts and serves to

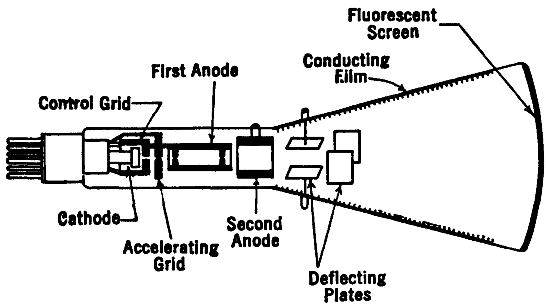


FIG. 37. Construction of a cathode-ray tube having electrostatic controls.

accelerate and concentrate the beam of electrons. The second anode has a still higher potential, usually in thousands of volts, and constitutes the final accelerating stage in the electron gun. After emerging from the second anode, the electron stream passes between the two pairs of deflecting plates. A difference of potential on one pair of plates will produce an electric field which will exert a force upon the electrons in the stream and deflect them in their path. Similarly, a difference of potential on the other pair of plates placed at right angles to the first will likewise deflect the electron stream. Thus through the medium of these two pairs of deflecting plates the electron beam can be directed upon any part of the fluorescent screen. The trigger for the electron gun is the control grid which receives an electrical input signal. The varying potential on this grid can start, stop, and control the intensity of the beam at all times. The electron beam impinging on the screen causes it to fluoresce in proportion to the intensity of the beam. Due to the persistence of vision of the eye, a moving beam produces a trace of light or, in the case of television, a complete image on the fluorescent screen. The fluorescent screen consists of a coating of a phosphor. Phosphors consist of compounds such as zinc silicate, cadmium tungstate, zinc sulphide, cadmium sulphide, and calcium tung-

state, or a mixture of these compounds together with some substance such as silver or copper. The choice of the phosphor is determined by the color and the persistence time desired. A conducting coating on the inside of the envelope of the tube serves to return the electrons from the fluorescent screen to one of the anodes. The secondary emission

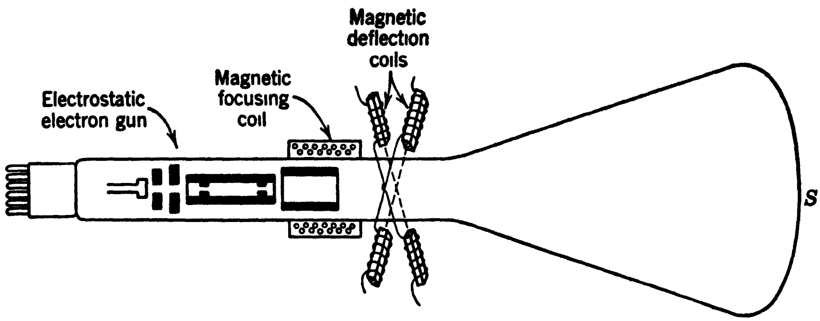


FIG. 38. Construction of a cathode-ray tube using magnetic deflection and focusing.

due to the impinging electrons removes electrons from the screen and prevents the building up of a negative charge.

Many cathode-ray tubes use magnetic fields for focusing and deflecting the electron beam. The construction of a tube of this type is shown schematically in Fig. 38. The electron beam is produced and accelerated by means of an electrostatic electron gun as explained in the preceding paragraph. The magnetic focusing is produced by a solenoid surrounding the tube. The coils for producing the fields for magnetic deflection are placed on the inside of a hollow cylinder like the stator coils in a-c machines. The magnetic focusing and deflection coils are usually formed into a single hollow cylindrical unit which slips over the cathode-ray tube. This unit is called a yoke. The general theory of electrostatic and magnetic focusing and deflection employed in cathode-ray tubes was covered under the subject of electron ballistics on pages 6–17.

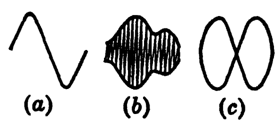
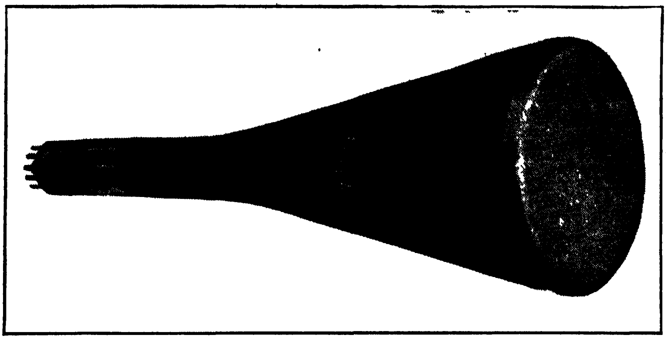


FIG. 39. Traces shown by an oscilloscope.

The principal applications of the cathode-ray tube are in oscilloscopes and in television receivers. In the oscilloscope a trace or curve is made on the screen which depicts the changes of voltage or current with time or some other variable. These traces (Fig. 39) can be viewed with the eye or photographed. In the television receiver a rapid



CATHODE-RAY TUBE

Heater, for unipotential cathode		
Voltage (a-c or d-c)		6.3 ± 10% volts
Current		0.6 amp
Direct interelectrode capacitances (approx)		
Grid No. 1 to all other electrodes		8.0 $\mu\mu$f
DJ1 to DJ2		1.3 $\mu\mu$f
DJ3 to DJ4		1.2 $\mu\mu$f
DJ1 to all other electrodes		9.5 $\mu\mu$f
DJ3 to all other electrodes		12.0 $\mu\mu$f
DJ1 to all other electrodes except DJ2		8.0 $\mu\mu$f
DJ2 to all other electrodes except DJ1		7.5 $\mu\mu$f
DJ3 to all other electrodes except DJ4		10.0 $\mu\mu$f
DJ4 to all other electrodes except DJ3		7.5 $\mu\mu$f
Phosphor		No. 1
Fluorescence		Green
Persistence		Medium
Focusing method		Electrostatic
Deflection method		Electrostatic
Maximum Ratings, Absolute Values		
Anode No. 2 and grid No. 2 voltage, max		2200 volts
Anode No. 1 voltage, max		1100 volts
Grid No. 1 (control electrode) voltage		
Negative value, max		125 volts
Positive value, max		0 volts
Peak voltage between anode No. 2 and any deflecting electrode, max		550 volts
Typical Operation		
Anode No. 2 and grid No. 2 voltage	1500	2000 volts
Anode No. 1 voltage for focus at 75% of grid No. 1 voltage for cutoff	337	450 volts
Grid No. 1 voltage for visual cutoff	−30	−40 volts
Deflection sensitivity		
DJ1 and DJ2	0.404	0.303 mm/d-c volts
DJ3 and DJ4	0.446	0.334 mm/d-c volts
Deflection factor		
DJ1 and DJ2	63	84 d-c volts/in
DJ3 and DJ4	57	76 d-c volts/in

FIG. 40. Cathode-ray tube. (Courtesy Radio Corporation of America.)

Average characteristics

E_f = 6.3 volts

Anode No. 1 volts adjusted to give focus

Curve	Electrode current	Anode No. 2 and grid No. 2 volts
A	Anode No. 1	2000
B	Anode No. 1	1500
C	Anode No. 2 and grid No. 2	2000
D	Anode No. 2 and grid No. 2	1500

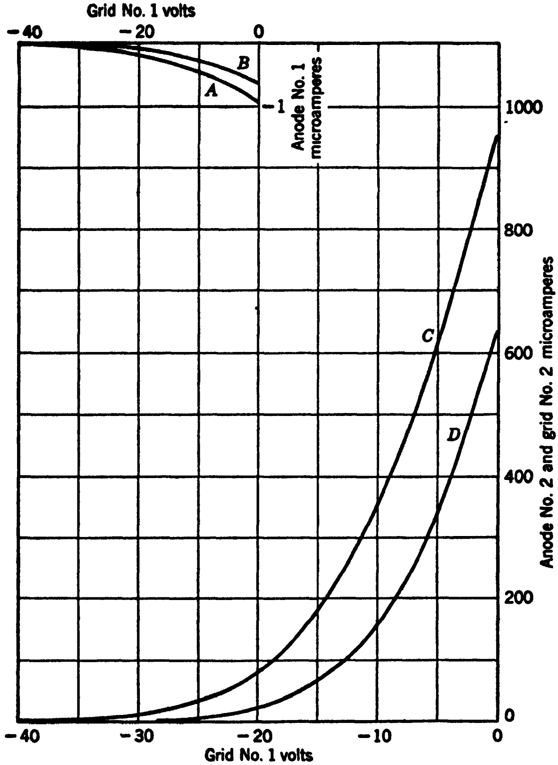


FIG. 40 (*continued*). Cathode-ray tube characteristics. (Courtesy Radio Corporation of America.)

scanning of the screen gives the impression of a complete image on the screen. The cathode-ray tube is very important as a measuring instrument for high-frequency changes. It holds this position because a beam of electrons or cathode rays passing from the cathode to the target possesses so little inertia that the beam can be deflected at a rapid rate, thus making possible the observance of high-frequency and transient phenomena. The power required to deflect the beam is negligible, and

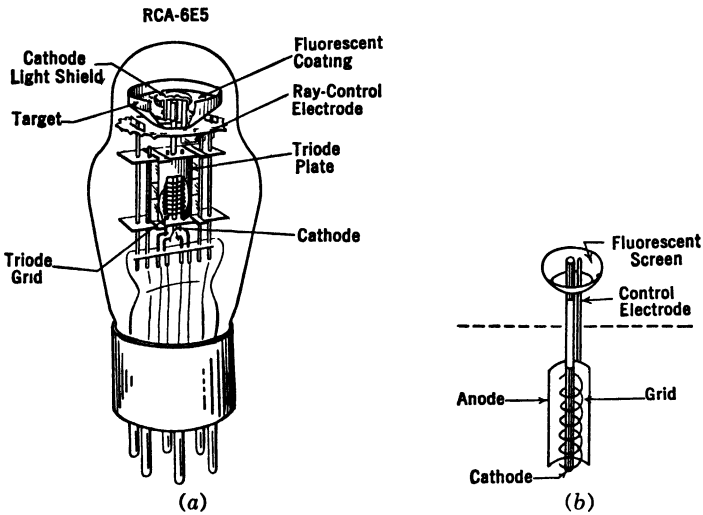


FIG. 41. Construction of an electron-ray tube. (Courtesy Radio Corporation of America.)

hence studies of weak signals can be made because so little power is required and the circuit conditions are not affected by the measuring equipment.

Nearly all cathode-ray tubes in use today are of the high-vacuum type. The functioning of these tubes is controlled by electrostatic and magnetic fields. Cathode-ray tubes are special tubes, but their special importance in the oscilloscope, which is used so widely in electrical measurements and tests, seemed to justify the introduction of their construction at this point. The rating, characteristics, and picture of a typical electrostatic cathode-ray tube are given in Fig. 40.

Electron-Ray Tube. The electron-ray tube is a combination of a triode and simple cathode-ray tube built into one envelope. Sectional views of this device are given in parts *a* and *b* of Fig. 41. The upper or dome-shaped part of the tube contains the cathode-ray part of the de-

vice, and the lower part houses the triode. A common heater-type cathode passes through the vertical axis of both sections, but the oxide coating covers the parts shown shaded only. In the upper section, the anode or target consists of a circular plate having a fluorescent coating on the upper side and a hole in its center. When a positive potential exists on this plate, electrons from the cathode are attracted to it and cause it to fluoresce with a green color. Without any other influence the circular plate would show a green circle of light with a black center. A grid called a ray-control electrode passes between the cathode and the inner circle of the fluorescent disk. This grid consists of a thin strip of metal placed parallel to the vertical cathode but insulated from both the cathode and the fluorescent disk. Now if this grid has the same potential as the disk, it will have little influence on the electrons going to the disk and a circle of green light still exists. However, if the grid becomes negative with respect to the disk, it will repel the electrons leaving the cathode and will cast a shadow on the circle as shown in Fig. 42. It should be noted that the grid or ray-control electrode for the top section is really an extension of the plate of the triode in the lower section. Hence the potential on the triode plate controls the potential of the grid of the top section.

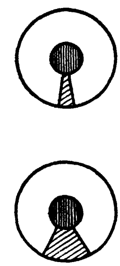


FIG. 42. Indications given by electron-ray tube.

The electron-ray tube sometimes carries the commercial name of the "magic eye" because of the resemblance to an eye suggested in Fig. 42. It is used as an indicator in sensitive electrical measuring devices and circuits, and in radio receivers to indicate the sharpness of the tuning for an incoming carrier signal. The sharpest tuning occurs when the shadow angle is at a minimum.

PROBLEMS

For the solution of problems requiring the use of characteristic curves of tubes, it is suggested that the curve sheet be covered with transparent paper held by short strips of masking tape. Light construction lines on this cover paper will not mar the original curves, and the overlay may be removed to serve as a part of the solution of the problem.

1. Determine the amplification factor of the triode in Fig. 11 for i_b values of 4 and 6 milliamperes, using the transfer characteristic.
2. Determine the amplification factor of the triode in Fig. 11 for i_b values of 8 and 10 milliamperes, using the average plate characteristic curves.

3. Determine g_m for the triode in Fig. 11 for $e_b = 200$ volts and i_b near 6 milliamperes; also for $e_b = 300$ with i_b near 3 milliamperes.

4. Calculate r_p for the triode in Fig. 11 for values of e_c at 0, -10, and -24 volts.

5. Calculate the amplification factor of the triode in Fig. 20 for e_c values near -1000, -1600, and -2000 volts.

6. Determine g_m for the triode of Fig. 20 at $e_p = 10{,}000$ volts.

7. Calculate r_p for the triode of Fig. 20 for $e_c = -1000$ and $e_b = 12{,}000$ volts.

8. Calculate r_p for the tetrode of Fig. 28 for $e_{c1} = 0$ and plate voltages of 20, 67, and 160 volts. Explain answer.

9. A tube has an amplification factor μ of 8 and an r_p of 9500 volts for a certain condition of operation. What is g_m in micromhos?

10. From the following data find approximate values of μ, r_p, and g_m at the point $e_b = 180$ volts, $e_c = -12.5$ volts.

e_b (volts)	e_c (volts)	i_b(ma)
180	-12.5	7.5
160	-10.0	7.5
180	-12.3	7.84

11. The pentode of Fig. 31 has a plate resistance of 7×10^5 ohms and a transconductance of 1575 micromhos. What is the value of mu for this condition of operation?

12. Anode No. 2 of the cathode-ray tube is operated at a potential 2000 volts higher than the cathode. With zero potentials applied to the deflection plates, what is the velocity of the electrons when they hit the fluorescent screen? If the deflection plates are parallel and extend 2 centimeters along the axis of the electron beam, what angle of deflection will be given to the beam by a uniform field of 50 volts per centimeter?

13. Assume that magnetic deflection coils are substituted in the preceding problem, producing a uniform field of 4.5 gauss for a distance of 4 centimeters along the axis of the tube. Calculate the angle of deflection for the electron beam after it has passed through this field (one pair of coils energized).

Chapter V

ELECTRICAL CONDUCTION IN GASES

For a long time gases were supposed to be perfect insulators. Dry air and other gases seemed to offer a high opposition to the flow of electric current. About 1900 J. J. Thomson performed experiments which showed that gaseous ions serve as carriers for the electric current.

The conduction of electric current in gases is not easily predictable since it depends on many variables. The resulting conduction may vary with the gas employed, the gas pressure, the potential between electrodes, the electrode material, the shape of electrodes, the distance between electrodes, the shape of the enclosing medium, and other factors. Conduction in gases may be attained with either hot or cold cathodes, but the action will be different in each case. All this is in contrast with the more simple theory of conduction of electricity in metals and in vacuums which has been considered in the preceding parts of this text. Some aspects of the phenomenon of gaseous conduction will be covered in this chapter to aid in the understanding of the action of gaseous electron tubes which are widely employed in the field of electronics.

Kinetic Theory of Gases. The molecules of a gas are in a constant state of motion similar to that of molecules in liquids and solids. Those in gases enjoy a greater freedom of movement so that what is termed gas pressure is really the result of multiple impacts of gas molecules upon the walls of the restraining enclosure. The simple kinetic theory of gases assumes the molecules to be small spheres which collide with each other in the course of their constant motion. The distance a molecule moves before it collides with another is called its free path. Obviously, the length of paths vary greatly, some being relatively short and others long. A study of the distribution of these paths will give a *mean length of free path* or average path which is of importance in the theory of electrical conduction in gases. The mean length of free path depends on the gas pressure and rises in magnitude as the pressure

falls and the molecules are farther apart. The mean length of free path will also depend upon the size of the molecules. Mean length of free path of gaseous molecules is of the order of 0.02 to 0.2 millimeter for a pressure of 1 millimeter of mercury.

An electron projected in a gas will likewise collide with the molecules present. After its first collision it will bounce off in a new direction until it experiences a second collision, and then in a new direction for a third collision, and so on. In this manner it will travel in a zigzag path through the gas. The distance between collisions is the length of free path for the electron, and the average length of these paths is the *mean length of free path* for the electron. Since the electron has a smaller mass and a smaller size than the molecules of gas, it will experience fewer collisions and will have a longer length of free path (approximately six times that of the molecule).

The molecules of the noble gases and mercury vapor consist of one atom. These atoms possess kinetic energy because of their thermal agitation, as suggested. In addition to this kinetic energy, these atoms may be given potential energy by displacements of electrons in their atomic structure. In the normal state the electrons in the atom exist in certain orbits, rings, or energy levels. If one or more electrons in an outer orbit are disturbed by the addition of energy, they may be moved out of their normal orbit or energy level into a higher level or they may even be removed from the atom. In this process the atom acquires new energy. This new energy may be released or radiated very quickly with a return to a normal state, or it may be retained for a short time before release. Thus a gas atom is capable of receiving, transporting, and releasing energy through rapid changes in its atomic structure. This property combined with its kinetic energy plays a very important part in electrical conduction in gases.

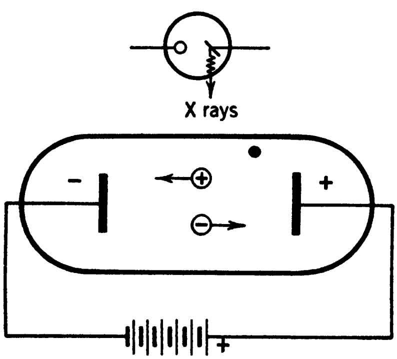


FIG. 1. Simple gaseous conduction.

Gaseous Conduction. The term gaseous conduction implies the conduction of current by a gas itself and within itself. Such conduction is simple to understand. The gas-filled tube of Fig. 1 contains two cold electrodes held at a difference of potential by a battery. Assume that one gas atom has been ionized to produce one positive

ion and one electron through the action of X rays. The electron will be attracted to the positive electrode and will enter the circuit leading to the battery. The positive ion will be attracted to the negative electrode and upon arrival will seize an electron from the plate and unite with it to form a neutral atom of gas. This simple dual action has removed one electron from the negative electrode and supplied one electron to the positive electrode. This transfer constitutes an electric current. Multiply this single transfer by millions and simple gaseous conduction results.

The ions taking part in gaseous conduction must be produced by some secondary action. Often such secondary action involves one or more forms of electron emission. When this happens the emitted electrons join the ion movement, and the total current through the gas consists of the pure gaseous conduction plus an electron current. Thus the electric conduction in gases usually involves two or more processes of electron transfer.

Methods of Producing Ions. Gases may be ionized (1) by thermal action, (2) by electromagnetic radiation, and (3) by collision with particles. Gases in a flame become ionized by thermal action. X rays, gamma rays, cosmic rays, and other forms of electromagnetic radiation such as ultraviolet light, have the power of ionizing a gas. Alpha particles and beta particles released from radioactive materials collide with gas particles and leave an ionized path. Much of our early and later scientific study has utilized the ionizing property of flying particles and the energy of radiation. In gaseous electronic devices, ionization by collision with electrons plays a very vital role.

Ionization by collision with electrons is produced in gases by the accelerating force of an electric field. The primary electrons for the ionizing process are generally released by thermionic emission or photoemission. To get a concept of ionization by collision, assume for the gaseous tube of Fig. 2 that the gas pressure and the electric field are of suitable value for ionization to occur. Electrons emitted from the negative hot cathode will be accelerated by the electric field toward the positive anode, resulting in collisions with atoms of gas while in flight. A certain collision with a Bohr atom shown in the figure results in ionization of that atom. Here the approaching primary electron is aimed at an electron in the outer orbit of the atom. This orbital electron is moving to the right, and the collision with the primary electron also moving to the right will increase its speed so that it will break away from the atom. The orbital electron carries a negative charge

$(-e)$ and the atom becomes a positive ion with a charge equal to $(+e)$ in magnitude. The primary impinging electron glances off at some angle as shown and continues its transit until it reaches the positive anode. During the ionizing collision the primary electron gives up a definite amount of energy to the atom. Hence the resulting positive ion acquires and carries a unit of energy which will be released again whenever the ion reverts to a normal atom.

Each electron passing between the electrodes of the tube of Fig. 2 under the influence of the electric field suffers many collisions with the atoms in the intervening space. Only a few (less than one per cent)

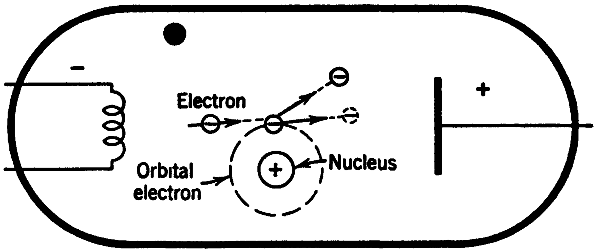


FIG. 2. Ionization by collision.

of these collisions result in ionization. This follows because the electron may not possess sufficient energy to produce ionization or because the nature of the collision is not favorable for ionization.

The preceding discussion might lead one to assume that the collisions of molecules and the impacts between flying electrons and atoms were actual physical contacts like those of a baseball bat hitting a ball. Such is probably not true. According to the concept of atomic structure, collisions and impacts may become the interaction of the charges on electrons, atomic nuclei, and molecules. As one molecule of gas approaches another, the electrons in their outer orbits exert a repulsion for each other, and when they approach close enough this force of repulsion may become sufficient to cause the molecules to stop and then bound away from each other. In like manner, when a high-speed electron approaches an atom, if the electron in the outer orbit happens to be in the direct path of the projected electron, the force of repulsion between the two may carry the one free from the atom and produce ionization. If one recalls that the ratio of the charge on the electron to its mass for moderate velocities is 1.76×10^{11} (mks system), it is obvious that the force exerted by the charges vastly outweighs the importance of the insignificant mass. It should be noted that, since atoms

are nebulous, electrons and even positive ions may pass through atoms provided they do not move on a line through the electrons or the nucleus.

The intensity of ionization in a gas depends on several factors such as applied potential, size of plates, distance between plates, and gas pressure. If two electrodes are placed in a gas at atmospheric pressure, a low difference of potential would sweep any ions present out of the gas. However, it would require a very high voltage to produce any ionization. This follows because the gas molecules are packed closely and any electron moving under the existing potential gradient could not be accelerated sufficiently before a collision to produce ionization. With a lowering of the gas pressure, the mean free path of the molecules increases, the electron moves farther, gaining more speed before collision, and ionization occurs more readily. The potential necessary to start ionization under any condition of gas pressure, etc., is often called the ionizing potential. As the pressure of the gas is reduced the ionizing potential will continue to decrease. Assume a moderate but fixed potential between the plates and then decrease the gas pressure by gradual steps. At a high gas pressure few if any electrons will ever attain sufficient speed to produce ionization. As the pressure is reduced the ionizing potential for that condition is reached and a current flows. A further reduction in pressure allows electrons to move farther before collision, and hence ionization becomes more intense. This increase of the intensity of ionization will continue with the decrease of pressure until the maximum number of the gas atoms is taking part in the ionization. Then any further decrease in gas pressure will reduce the number of atoms present so that the actual current conducted by gas ions will decrease progressively with pressure, ultimately approaching zero in a very high vacuum.

The term ionizing potential is frequently used rather loosely. As commonly applied, it refers to that potential which must be applied between two electrodes to produce ionization for a given gas, pressure, temperature, and electrode spacing. In a specific sense, one may think of the ionizing potential as the voltage through which an electron must fall to produce ionization. Following this concept, each gaseous element requires a definite energy in electron-volts to produce excitation and ionization. For example, helium has a minimum excitation potential of approximately 20 volts and an ionizing potential of approximately 25 volts. For mercury vapor the approximate excitation value is 4.7 volts with an ionizing value of 10.4 volts. The distance through which an electron can fall before collision depends upon the

gas pressure and temperature. Obviously, these latter factors govern the potential gradient and the difference of potential between electrodes for producing ionization.

The collision of the electrons with gas atoms may be divided into four groups. First, many collisions result in a mere "bouncing off" or change of direction of the electron where little or no energy is imparted to the atom. These are known as *elastic collisions* and constitute the majority of all impacts. Second, some collisions occur where the energy of the electron is insufficient to produce ionization but is great enough to move an outer electron out of its orbit or energy level. Here the atom has been given some energy and is said to be *excited.* This excited state of the atom is a very transient one and the atom releases the acquired energy almost immediately. The energy is released in the form of electromagnetic waves. This wave energy radiation may be *visible light* and as such is very important in some types of electric-lighting sources. In a third group of collisions, the electron is able to impart energy to the atom so that it is retained for a short period of time. While this energy is retained, the atom is said to be in a metastable state. A metastable atom may receive additional energy from a second collision which will be enough to produce ionization. It should be noted that the excited atom and the metastable atom retain all electrons and hence do not carry a charge and are not affected by an electric field. The fourth group of collisions is the *ionizing collision.* If the conditions are favorable for ionization to occur, all four groups of collision will be taking place.

Gases may be ionized by the collision of positive ions with atoms. However, the positive ions present in gaseous tubes are not very effective ionizing agents. Their large mass relative to the electron prevents rapid acceleration in an electric field and their larger size results in more collisions with molecules, so they have little opportunity to acquire velocities sufficient for ionizing collisions. It is well to remember that in the simplest gas, hydrogen, the positive ion (a proton) is 1840 times as heavy as the electron. In the more complex gaseous elements used in commercial tubes, the positive ion has a mass several thousand times as large as the electron (see Table 1, page 7).

Deionization. Deionization is the reverse process of ionization. It is effected by a recombination of positive ions and electrons to form neutral normal atoms. Such recombinations are accompanied by a release of the energy required for ionization and imprisoned in the positive ion during its existence. Recombination may take place (1) inside the volume of gas, (2) along the walls of the gas-enclosing

chamber, or (3) at an electrode under the attraction of an electric field. Recombinations do not take place readily inside the gas volume if an electric field is present because positive ions and electrons will have high relative velocities in opposite directions. If the conditions are favorable for the formation of negative ions (atom plus an electron), recombination may take place readily since the negative ion with larger mass moves more slowly than an electron, even in the presence of an electric field. Also the negative ion gives up its extra electron readily to a positive ion. Positive ions which diffuse to surface walls obtain electrons readily for deionization. Electrodes having negative potentials attract positive ions and supply electrons for recombinations. The energy released during deionization may be in the form of heat and may aid electron emission at the cathode. The preceding statements on recombinations apply to low-pressure discharge conditions but not to high-pressure discharges where mean free paths are negligible compared to electrode spacing.

Stages of Electrical Conduction in Gases. Electrical conduction in a gas between cold electrodes may pass through and exist in three successive stages. These stages are called the Townsend discharge, the glow discharge, and the arc. The term discharge arose from early experiments wherein gaseous conduction served to discharge electricity stored in condensers. The Townsend discharge can be explained by the circuit and curve of Fig. 3. Two cold plates are placed in the chamber C containing gas under reduced pressure. A variable voltage from source E_x is placed across the plates. Beginning with X at zero or a, the voltage is raised as X moves to position b with a current flow in the circuit as shown by the curve $OPQR$. While no ionizing agent is provided within the chamber, a small number of ions will be present because of cosmic rays and other electromagnetic radiations which exist outside of the chamber or because of radioactive substances which may be present on the inside or outside. The small number of ions present and being formed constantly will be attracted to the electrodes and will constitute a *very minute current* which will rise along the line OP under the influence of a weak field. The voltage at P is sufficient to sweep all these ions out of space as rapidly as they are formed, and from point P to Q there will be no change of current. At point Q the ionizing potential for the pressure and gas used is reached. Hence a few ions will be produced and their transfer will add to the very minute current present. Now, as the voltage is raised above the ionizing potential, more primary electrons will produce ionization.

Also, as the potential rises, two things may happen. First, an occasional primary electron may have more than one ionizing collision while in transit, and, second, a few electrons formed by collisions may in turn effect an ionizing collision. In this manner, the process of ionization becomes accumulative and the gaseous conduction current rises rapidly following an exponential type of curve. Ultimately, as the point X is moved to the right, the current tends to rise without

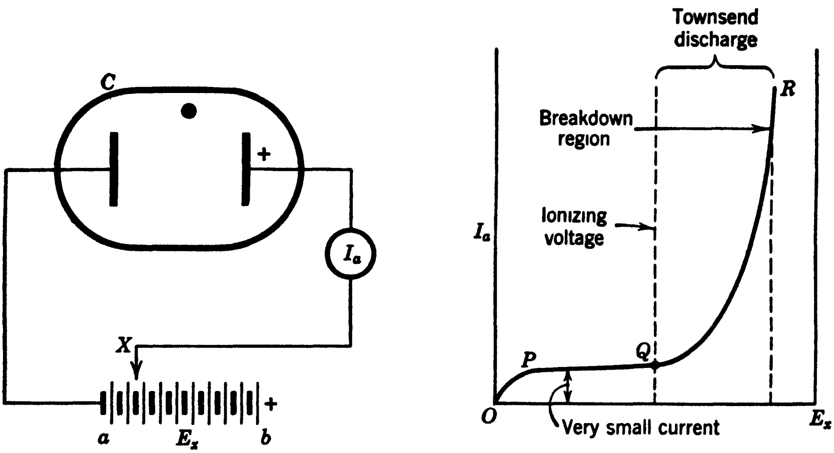


FIG. 3. Circuit and phenomena for Townsend discharge.

apparent limit until a breakdown region is reached. The breakdown region is marked by two occurrences. First, the operation of the tube becomes self-sustaining, and, second, the discharge passes into a second stage or form of discharge. Up to this point the discharge has depended upon the production of some ions by an external or separate agent, such as electromagnetic radiations. If this external action had been stopped, the discharge current would have stopped likewise. At the breakdown point a stage has been reached where the products of ionization (electrons and positive ions) will reproduce sufficient electrons and ions to keep the discharge continuous. This action is accomplished by electron emission at the cathode resulting from bombardment by the positive ions attracted to it plus the energy released during recombination. One factor in the breakdown of the Townsend discharge is the change in the potential distribution between the electrodes resulting from the rising positive space charge which increases the potential gradient near the cathode. After reaching the breakdown point, the Townsend discharge may pass into other types of discharge

such as the corona, the glow, and the arc. The term breakdown potential is known also as sparking potential and ignition potential.

The breakdown voltage E_b in the Townsend discharge depends upon the geometry of the electrodes, their work function, the gas employed, the gas pressure, and the distance between the electrodes. In order to generalize on the phenomenon, a particular gas and electrode material may be chosen and the cold electrodes may be assumed to be parallel plates, thus giving a uniform electric field if space-charge effects are ignored. If the electrodes are held at a fixed spacing and the gas pressure is increased from zero to X, the trend of the breakdown voltage E_b varies as shown in Fig. 4. Here the mean free path of the electron will control the intensity of ionization and, in turn, the breakdown point, as previously discussed on page 99. Next, assume the gas pressure is held constant and vary the spacing between the electrodes from zero to some considerable distance X. The measured breakdown voltage will again follow the same trend, as shown in Fig. 4. Here also, the length of free path of the electron determines the characteristic curve. At short spacings the electrons experience few collisions and the necessary breakdown voltage must be very high. For large spacing the distance may equal several lengths of free path, and again E_b must be high. The optimum condition for breakdown is likely to occur at a spacing approximating the mean free length of path for the conditions assumed. This characteristic variation of breakdown voltage with electrode spacing has been utilized in the design of cold-cathode rectifiers and grid-glow tubes to be covered in the following chapter. Thus a spacing corresponding to point b on Fig. 4 requires a relatively low breakdown voltage, whereas a short spacing such as point a will permit three times such potential before conduction occurs.

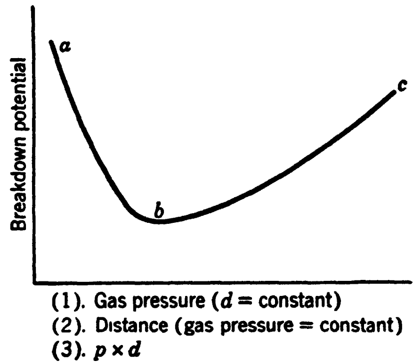


FIG. 4

The preceding discussion discloses that the trend of change of breakdown voltage with variation of gas pressure for a constant electrode spacing is the same as for the inverse condition, namely, a variation of electrode spacing for a constant gas pressure. From this disclosure it is obvious that a similar trend will follow for a variation in the

product of gas pressure times distance, or $p \times d$. This relationship is known as *Paschen's law.* Paschen's law states that, for parallel plane electrodes in a given gas at a given temperature, the breakdown voltage is a function of the product of the pressure and the electrode separation. To understand this law one should remember that under the conditions assumed the mean free path is inversely proportional to the pressure, so that the ratio between the spacing and the mean free path remains

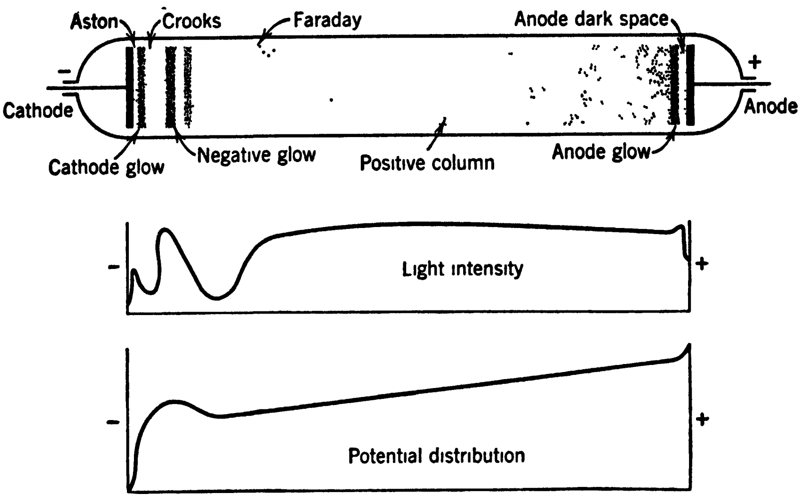


FIG. 5. Distribution of quantities in a long gaseous column with a cold cathode. (Reprinted with permission from L B Loeb, *Fundamental Processes of Electrical Discharge in Gases,* John Wiley & Sons, 1939, p. 566, Fig. 269)

the same. Thus the number of free paths between the electrodes remains the same for a particular product of $p \times d$.

The *glow discharge* is a self-maintaining discharge which is marked by a luminous column, a low current density, and a high voltage drop between electrodes. This discharge gets its name from its soft luminous effect exemplified in the familiar neon sign. Some of the properties of the glow discharge for a long tubular chamber under an applied d-c potential are indicated in Fig. 5. Most of the light comes from the long positive column with supplementary bright areas at the cathode glow, the anode glow, and the negative glow points. The potential distribution shows a rapid rise of voltage close to cathode, then a gradual rise throughout the long positive column, with a final spurt or rise in potential near the anode. This form of potential distribution may be explained by the slow mobility of the positive ions in the long positive

column. The ions form a positive space charge which moves down to the cathode and gives a form of virtual anode close to the cathode. The rapid rise in voltage beginning at the cathode is equal to the breakdown voltage for the particular pressure and gas used. This voltage is several times the ionizing potential and it produces primary electrons and ions in sufficient numbers to make the discharge self-maintaining. This rise in voltage, often called the cathode fall of potential, depends upon the material used in the electrodes, upon the kind of gas, and upon the gas pressure in the tube. This voltage lies within the range of 50 to 300 volts. The positive ions are repelled by the positive potential at the anode so that only electrons exist in a narrow region called the anode dark space. In this region a more rapid rise of potential (stronger electric field) is needed to withdraw the electrons from the positive column.

When the current in the glow discharge is small the cathode glow covers a small area of the cathode, and the area increases in size with the magnitude of the current. Since the tube voltage is nearly constant with current variation, it is necessary to limit or control the current by some external means. In neon signs this control is built into the transformer that supplies the voltage. The constant voltage characteristic of the glow discharge is utilized in voltage-regulator tubes to be described later.

If the current density in the glow discharge is permitted to pass a certain maximum value for a particular gas, the character of the discharge changes rapidly. The increase of current will be accompanied by a rise of voltage drop between the cathode and anode and the glow will rise in intensity. This new state is called the abnormal glow. A further increase of current causes an acceleration in the brightness of the glow and a rise in voltage until the discharge suddenly changes to a new stage called the *arc*. The important characteristic of the arc is shown in Fig. 6. Here the voltage across the arc varies inversely with the current through the arc. This is a negative resistance characteristic which tends to damage any equipment associated with the arc. Hence the arc calls for a stabilizing element in the external circuit which will limit the current.

The phenomenon of the electric arc discharge is not well understood. The arc itself is made up of positive ions, electrons, excited atoms, and metastable atoms. The high current density in the arc results in very high temperatures. This high temperature may influence the arc phenomenon in several ways. First, the high temperature may increase electron emission from the cathode and thus increase the current.

Second, the temperature may be sufficient to produce ionization by thermal agitation. Third, the high temperature reduces the ionizing energy necessary to produce ionization, thus permitting an increase of ionization by collision. All these processes may act accumulatively to produce intense electrical conduction even with exceedingly short electrode spacing. If the arc takes place in the presence of oxygen the electrodes will be burned.

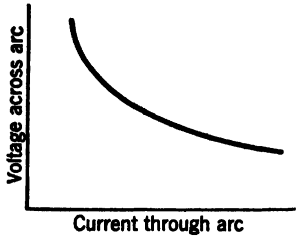


FIG. 6. Characteristic of the electric arc.

The three stages of gaseous discharge with cold cathodes are briefly summarized in Fig. 7. In the Townsend discharge initial conduction is produced by some ionizing agent and a minute current exists under a rising potential until the ionizing potential is reached. After reaching this point the conduction current rises rapidly with voltage, and the potential distribution shifts to bring about a new stage wherein the discharge is self-sustaining. In the new stage (the glow discharge) the potential drop across the tube is constant while the current rises until the cathode glow covers the cathode surface. A further rise in current flow results in a rise in cathode-anode drop, a rise in luminous brightness, and a change to the arc discharge. The arc discharge is marked by an intense luminosity and a fall of potential with a rise in current. Thus the transitions in the three stages are marked by (1) a rise of current and a rise in potential, (2) a rise of current and a constant potential, and (3) a rise of current and a fall of potential.

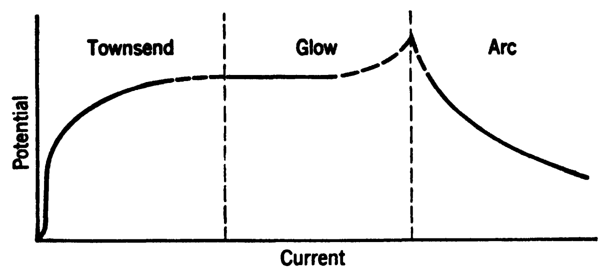


FIG. 7. Summary of the stages of discharge between cold electrodes in a low-pressure gas.

In all forms of electrical conduction in gases the positive ions play the vital part. The large mass and slow mobility of these ions cause them to form a dense concentration in the space where conduction is

taking place. Langmuir has termed such concentrations or positive space-charge regions the *plasma.* The plasma has boundaries on all sides. For example, in Fig. 5, the end boundaries occur at the negative glow and the anode glow. The other boundary is the outside of the column. This outer boundary may be the wall of the tube but in general the outside boundary is a layer or sheath. This *sheath* is a region consisting of ions, metastable atoms, and normal atoms. The sheath differs from the plasma in that the ions are in a general state of diffusion and do not have a coordinated drift as a part of the current. The sheath can be likened to a thin layer of water on the inside of a pipe having a rough inner surface. This outer skin of water may be nearly stationary and yet serve as a boundary for the water flowing through the pipes. In gaseous conduction tubes, sheaths may represent the boundaries for electrodes as well as enclosing surfaces.

The preceding discussion of electric conduction in gases has assumed the use of a cold cathode, whereas most of the gaseous tubes in commercial use employ hot cathodes. The hot cathode furnishes a copious supply of electrons continuously. These electrons add greatly to the conduction of current (electron flow) and reduce the voltage drop required for producing simple ionization. The conduction of current is non-self-sustaining and the characteristics differ greatly from conduction in vacuum or self-sustaining conduction in gases. The characteristics of the hot-cathode gaseous tube will be covered in the chapter which follows.

Gaseous Lighting Units. Many lighting units of both novel and useful design operate on the principle explained in the preceding articles. One of the first of these is the familiar Geissler tube which usually consists of a long glass tube bent into irregular shape and containing a number of sections, each having a different kind of gas under low pressure. The sections are connected by conductors and the whole tube is placed in operation by the application of a high-voltage alternating current to the terminals. Very striking color effects are obtained when the tube is placed in operation and viewed in the dark.

The neon light consists of a long glass tube about ½ inch in diameter and filled with neon gas under low pressure. It is a very efficient source of light, giving a reddish-yellow color. The low cost and high efficiency of this light source has resulted in its popularity for sign lighting all over the world. Argon gas and sometimes helium and mixtures of gases are used to give different colors in sign lighting.

Glow lamps consist of bulbs filled with an inert gas such as neon under low pressure and containing two electrodes. These lamps use

the principle of ionization and the accompanying production of light arising from negative glow for illumination or for voltage regulation. One form of glow lamp is placed in an incandescent-type lamp bulb having a high resistance in its base. The resistance limits the magnitude of the current of gaseous conduction. These lamps consume from $1/25$ to 5 watts and are used for night lamps, signal lamps, and Christmas-tree lamps. Several lamps of this type are illustrated in Fig. 8.

The Cooper-Hewitt mercury-vapor lamp consists of a long tube containing a small quantity of mercury. Ionization of the mercury is

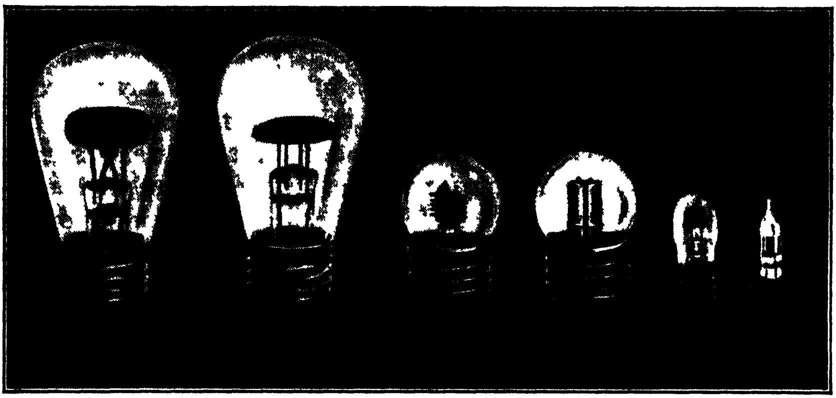


FIG. 8. Neon glow lamps (Courtesy General Electric Company.)

started by tilting the tube, by applying a high inductive voltage, or by the use of thermionic emission from filaments. The color of the light, bluish yellow and devoid of red rays, causes human faces to have a deathly pallor. The specific advantages of this light are that it casts but little shadow, covers a narrow band in the visible light spectrum, and gives excellent visual acuity. Formerly the Cooper-Hewitt unit had fairly extensive application for drafting rooms, machine shops, and industrial processes where detection of color was not essential but where fine definition was important.

Today this low-pressure type of mercury-vapor lamp is being replaced by the fluorescent lamp which provides a lower cost and higher efficiency. Moderate- and high-pressure mercury-vapor lamps are concentrated and efficient sources of light that have many special applications.

The fluorescent lamp consists of a long glass tube coated on its inner surface with a fluorescent material known as phosphor and containing a small amount of mercury vapor. Emitter electrodes at the ends of the

tube produce initial electrons for ionizing the mercury vapor and furnishing ultraviolet light. The ultraviolet light causes the phosphor to fluoresce and give forth visible light. The different phosphors available are capable of producing different colors of light, including white, at relatively high efficiencies compared to the heated-filament type of light source.

The gaseous conduction of mercury vapor has given rise to a number of important developments. Much of the electromagnetic radiation from the ionized mercury vapor is ultraviolet light. Ultraviolet rays have a high therapeutic and sterilizing value. Accordingly, one type of mercury-vapor lamp (called a sun lamp) is built to give healthful light for man and animals. A special form of this lamp (germicidal) is used for sterilization in refrigerators, in operating rooms, and in treatment of wounds.

The sodium-vapor lamp uses the vapor of sodium as the gaseous conducting medium. A small quantity of metallic sodium is contained in an inner chamber. It requires several minutes for the lamp to heat up, evaporate the sodium, and come to full normal brilliancy. The lamps give a yellow light at a very high luminous efficiency and are used for highway lighting.

Chapter VI

GASEOUS AND VAPOR ELECTRON TUBES

Gaseous versus Vacuum Tubes. Two criteria for distinguishing between vacuum and gaseous electron tubes are (1) the pressure of the gas and (2) the mean free path of the electrons. In a highly evacuated tube of 10^{-8}-millimeter pressure at 0 degrees C, there are more than ten billion molecules in each cubic centimeter. For the ordinary vacuum tube of 10^{-6}-millimeter pressure, the mean free path of the electron is of the order of 42 meters. Since the electrode spacing in vacuum tubes never exceeds 2 or 3 centimeters (except in cathode-ray tubes), it is obvious that electrons move freely from cathode to anode with rare collisions with molecules of gas. In contrast with this picture, in gaseous tubes the normal spacing distances of electrodes is much greater (up to several inches), whereas the mean free path for common gas pressures of 10^{-2} to 10^{-1} millimeter is of the order of *a millimeter* or less. Thus in gaseous tubes the electrons suffer many collisions in transit and are likely to produce an intense ionization of the gas.

Neutralization of Negative Space Charge. Positive ions formed in an electron tube tend to neutralize the negative space charge surrounding a hot cathode. If a small amount of an inert gas is admitted to a high-vacuum diode, the gas molecules will be ionized by collisions with the emitted electrons in transit to the anode. The electrons formed by collision will be attracted to the positive anode and, because of their small mass, will reach it quickly. The positive ions formed by collision will move to the negative cathode rather slowly because of their large mass. Since the positive ions move slowly, they will remain in the cathode-anode space for a long time relative to the time of transit of electrons. Accordingly, a large number of positive ions may exist at a given instant within the region of the negative space charge surrounding the hot cathode. A single positive ion close to an electron will *neutralize the charge* of that electron at that instant. Thus the field produced by this pair is zero (neutral space charge). In an infinitesimal fraction of a second the electron of the pair is whisked away

by the electric field and the positive ion moves a short distance toward the cathode. In its new position the positive ion will be close to a second electron and at that instant will serve to neutralize this electron. In this manner the slow mobility of the positive ion will permit it to neutralize many (perhaps hundreds) of electrons while passing through the negative space-charge region. It is readily conceivable that, if at every instant the region surrounding the cathode contains the same number and distribution of positive ions and electrons, the negative space charge will be completely neutralized. This ideal state may seldom exist, but any degree of neutralization will overcome partially the negative space charge and will increase the number of primary electrons that are attracted to the anode. In passing, it should be noted that, if the positive ions present close to the cathode exceed the number of electrons, they will act like a positive grid very close to the cathode and thereby will increase greatly the primary electron flow to the anode.

The conduction of electricity in the gaseous hot-cathode tube may be considered as consisting of the three following components:

1. The primary electron current existing in a high vacuum.
2. The gaseous conduction resulting from electron transfer by ions.
3. *The increased flow of electrons of thermionic emission resulting from the neutralization of negative space charge.*

The third component is the largest in magnitude and of the greatest importance. Component number two is usually small in magnitude. The addition of a small amount of gas or vapor greatly increases the current rectified by a diode. It also reduces the voltage required between the cathode and anode. This action is illustrated in the curves in Fig. 1*a*, giving a comparison of the anode-current anode-voltage characteristics of hot-cathode vacuum and gaseous tubes. In the gaseous tube with negative space charge neutralized, the anode current rises abruptly after the ionizing potential is reached, giving a constant anode-cathode drop with load. Thus for the gaseous tube the rectified power output is increased and at the same time the power loss within the tube is reduced. Both these changes increase the efficiency of rectification of alternating current.

Potential Distribution in Hot-Cathode Gaseous Tube. The presence of positive ions in gaseous and vapor tubes has a marked effect upon the potential distribution between the cathode and anode. In a high-vacuum tube the negative space charge depresses the voltage near the cathode, whereas in the gaseous tube the positive ions tend to neutralize

this negative space-charge effect. Furthermore, if sufficient gas or vapor is present, an arc form of discharge takes place. The positive ions form a plasma for the arc which extends close to the cathode giving the potential distribution as shown in Fig. 1*b*. The trend of this curve is similar to that of the glow-discharge tube but here the magnitude of the cathode drop from plasma to cathode is much lower. In the glow-discharge tube a high voltage is necessary to produce electrons from a cold cathode, whereas in the tube under consideration a copious supply of electrons is emitted thermionically. In the gaseous and

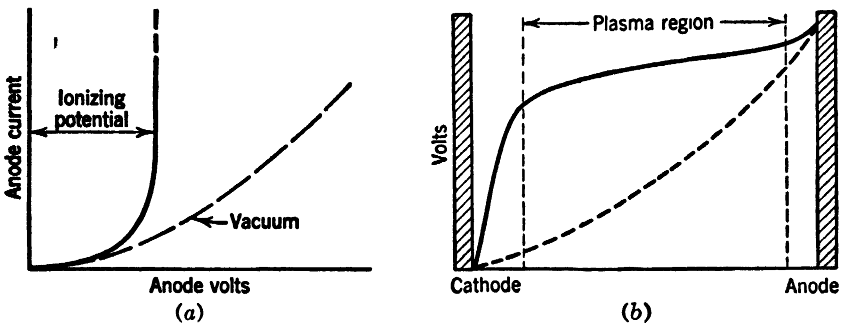


FIG. 1. Current and potential characteristics in a hot-cathode gaseous or vapor-rectifier tube.

vapor-arc tubes the total fall of potential from anode to cathode is of the order of 10 to 25 volts. The exact form of the potential distribution curve will vary with the gas or vapor and the pressure used and somewhat with the geometry of the tube. In general, there will be a rise of potential from the plasma of the arc to the anode for extracting the electrons out of the plasma. The potential distribution which would exist in vacuums because of negative space charge is indicated by the dotted line in Fig. 1*b*.

Cathode Sputtering. The positive ions produced in a gaseous tube bombard the hot cathode. When the ions hit the cathode they give up energy in two ways. First, the positive ions possess kinetic energy arising from the velocity acquired in the electric field. Secondly, the positive ion carries potential energy resulting from ionization. This latter energy is released when recombination takes place at the cathode. Both forms of energy given up when the positive ion hits the cathode may be transformed into heat, thus raising the temperature of the cathode and increasing the rate of thermionic emission. This process is utilized in the ionic-heated cathode and will be referred to later. If

the electric field is too high the bombardment of the positive ions may disintegrate the emitting surface. This action is known as *cathode sputtering* and the removed active material such as thorium or barium may land on a nearby electrode such as a grid and result in emission from the electrode. The loss of active material from the sputtered cathode will reduce the emission from the cathode and may make the cathode inoperative, thus ruining the tube for further service. Cathode sputtering results from (1) insufficient cathode emission (underheating), (2) too large cathode-anode current (overload), or (3) too high cathode-anode voltage drop. These conditions have a direct interrelation so that two of them occur simultaneously. Cathode sputtering can be prevented with a suitable resistor in the load circuit to limit the current output of the tube to the rated value and by having the cathode at normal emission temperature and the correct bulb temperature before anode voltage is applied.

While cathode sputtering is very harmful in an electron tube, the sputtering process may be useful in some manufacturing processes. Thus a thin coating of a metal may be placed on a plate or an electrode by sputtering this metal from a second electrode serving as a cathode.

Inverse Peak Voltage. The applications of any electron tube having unilateral conductivity depends largely upon the maximum inverse or reverse peak voltage that may be applied in the cathode-anode circuit without a reverse current flow. The term inverse voltage does not refer to the insulation or dielectric strength of the tube parts. In the high-vacuum tubes previously discussed it was pointed out that, with a cold anode, exceedingly high inverse voltage (anode negative and cathode positive) could be employed without any danger of a reverse current (called breakdown or arc-back). However, when gas is present in a tube conduction may take place from a negative cold plate, as was explained in Chapter V, for the Townsend discharge, the glow discharge, or the arc discharge. Accordingly, in gaseous and vapor electron tubes the maximum inverse voltages must be held at much lower values.

The inverse voltage across a rectifying tube without a filter is the maximum value of the impressed a-c voltage. If a tube supplies a load through a filter, the maximum inverse voltage between the cathode and anode will be equal to the impressed a-c maximum voltage plus the d-c voltage at the input to the filter. This relation is illustrated for the half-wave rectifier and filter circuit shown in Fig. 2. The magnitude of the inverse voltage present will depend upon the filter and load circuit used and may approach a value twice the crest of the input.

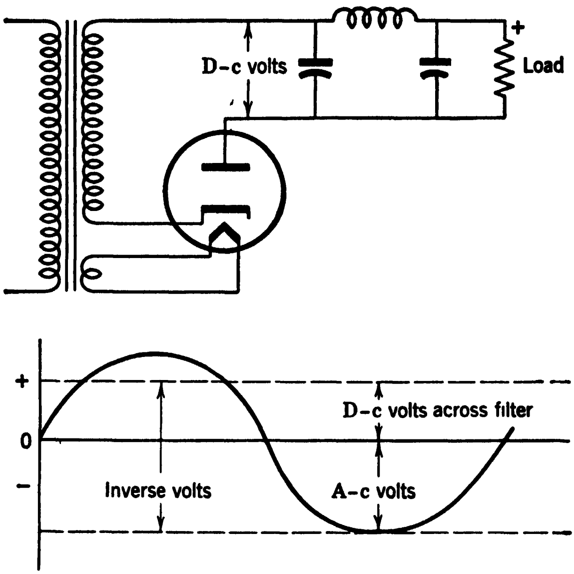


FIG. 2. Inverse voltage across a vacuum tube supplying a load through a filter.

Effect of Gas Pressure. Inert gases used in incandescent lamps reduce the evaporation of tungsten and permit higher operating temperatures and higher luminous efficiency for a given life. In a similar way, the use of inert gas in an electron tube reduces the evaporation from the cathode and permits the use of higher cathode temperature with corresponding increase of emission current. The higher the gas pressure, the greater the permissible thermionic emission.

The inverse peak voltage of a gaseous or vapor tube decreases rather rapidly with a rise in pressure. As the gas pressure rises above zero, the ionization by collision increases. The tendency to arc-back depends upon the number of positive ions present, and hence the permissible gas pressure will be determined by the inverse peak voltage required. In tubes employing mercury vapor the temperature of the tube becomes an important factor because the amount of mercury evaporated, and hence the vapor pressure, depends upon the coolest place in the tube where the vapor condenses. Thus in a mercury-vapor tube the operating temperature may become very critical as far as peak inverse voltage is concerned.

Gaseous Rectifier Diodes. The commercial gaseous rectifier diode is filled with argon gas under a pressure of approximately 2 pounds. The argon gas furnishes the positive ions for neutralizing negative

space charge and the high gas pressure reduces evaporation and thorium loss from the cathode. The anode is a graphite disk with its lead brought in from the top of the tube. The cathode is a helical coil of tungsten or thoriated-tungsten wire connected through leads to a screw type of base, as illustrated in Figs. 3 and 4. The cathode filament is operated at a low voltage of 1.8 to 2.2 volts and with currents of 6 to 27 amperes, depending upon output current. The tube "fires" or conducts with a nearly instantaneous start. No harm results from impressing voltage on the anode before the cathode reaches operating temperature because the relatively high gas pressure limits the velocity (collisions) of the positive ions. The average d-c pick-up voltage for the tube is 11 to 13 volts and the average arc drop (cathode to

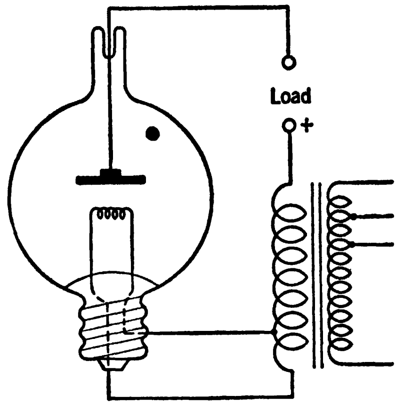


FIG. 3. Circuit for a half-wave Tungar rectifier.

Anode	
Load amperes (d-c)	6
D-c pick-up volts	12
Arc voltage	7
Maximum inverse peak	300
Filament	
Volts	2 2
Amperes	18

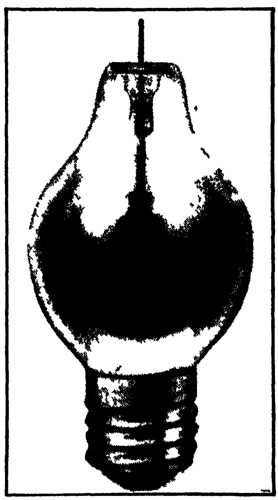


FIG. 4. Typical gaseous rectifier tube. (Courtesy General Electric Company.)

anode) is only 7 volts. The maximum inverse peak voltage varies from 150 for small-capacity tubes to 300 for the 6-ampere capacity. This low inverse peak is necessary to prevent a glow discharge from

starting on the nonconducting half of the cycle of rectification.

These commercial gaseous rectifier tubes are known under the trade names of Tungar and Rectigon. Small tubes for trickle charging have an output capacity of 0.5 ampere. Larger tubes have output capacities of 2, 6, and 15 amperes. The usual applications for these tubes has been charging storage batteries and supplying direct current for arc lamps in motion-picture projectors. Complete charging units are designed for service in the home or commercial garage to charge from one to six automobile storage batteries in series. The circuit for a simple unit is given in Fig. 3.

Mercury-Vapor Rectifier Diode. The mercury-vapor rectifier diode uses a hot cathode and mercury vapor under a *low pressure.* It should not be confused with the pool type of tube which uses a pool of cold metallic mercury for a cathode. A small quantity of mercury is inserted in a hot-cathode evacuated tube. A part of or all the mercury vaporizes and the vapor atoms ionize and serve as the conducting medium in the tube. The conduction is of the arc type. In comparison with the vacuum diode the mercury-vapor tube carries a much larger current, has a neutralized negative space charge, and has a nearly constant cathode-anode voltage drop (within the range of 10 to 20 volts) which is nearly independent of current. The absence of negative space charge and its accompanying losses allows a larger electrode spacing and smaller-size electrodes for a given current-carrying capacity. It also permits the use of an electron-emitting cathode of higher efficiency and much larger current-carrying capacity. The anodes consist of disks made of metal or graphite which are mounted on leads brought out at the top of the tube. The hot cathodes are of the oxide-coated type and are generally placed within a hollow metal cylinder to reduce radiation losses. Various configurations have been used in the construction of these cathodes as illustrated in Fig. 5.

Glass-enclosing tubes are used for mercury-vapor units having small current-carrying capacities. Metal-enclosing tubes are used for the larger capacities.

The mercury-vapor rectifier is superior to the gaseous rectifier in that it will withstand much higher inverse peak voltages of the order of 5,000 to 10,000 volts. However, the mercury-vapor diode has the disadvantage that voltage should not be applied to the anode until the cathode and the entire tube reach normal operating temperature. Such application of voltage when the cathode is cold or partially heated will result in the destruction of the cathode-emitting surface by positive ion bombardment. Special care in the installation and operation of mer-

cury-vapor tubes is necessary to insure long life and satisfactory operation. The user should observe carefully the manufacturer's instructions for operation in an upright position, long warm-up period when

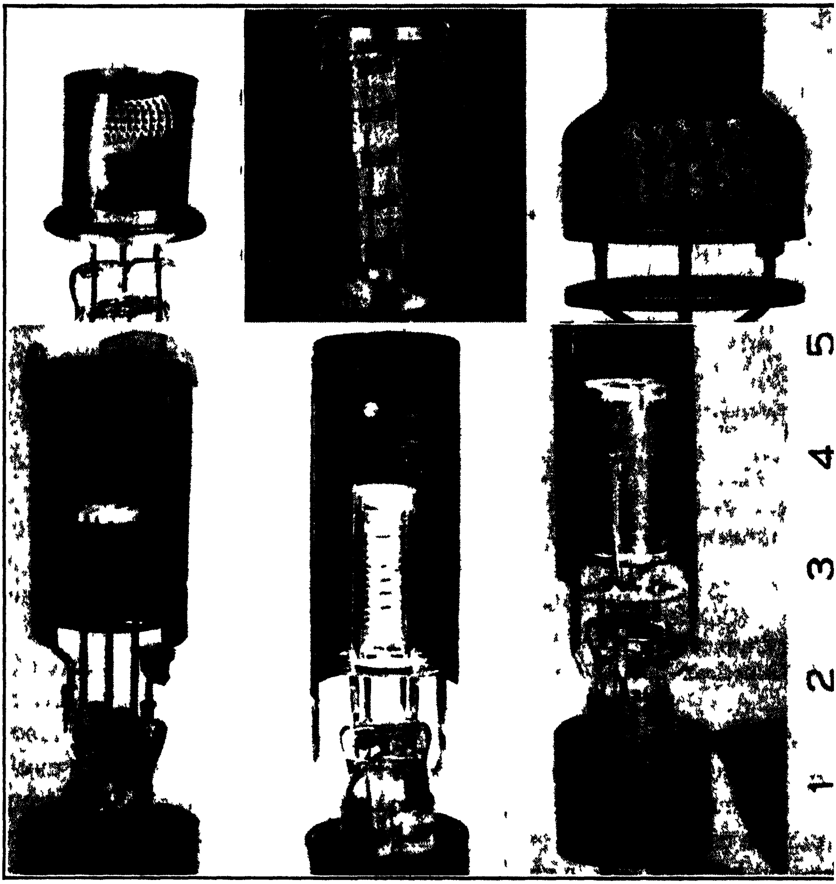


Fig 5 Assemblies of oxide-coated cathodes used in gaseous and vapor tubes. (Courtesy General Electric Company)

first installed, suitable surrounding temperature, protection from water drops, acid fumes, and so forth.

The accepted name for the mercury-vapor diode in the power industry is the *phanotron.* These tubes have average current output ratings of from 0.25 to 30 amperes. A tube of low rating is shown in Fig. 6 and one of high current in a metal enclosure in Fig. 7. Phanotrons are used for supplying direct current for applications where medium values of voltage and current are required.

Some diodes termed Tungars contain mercury vapor and operate at a higher pressure comparable to those containing inert gas. These mercury-vapor tubes possess characteristics similar to those containing gas. Similarly some diodes containing inert gas operate at pressures comparable to phanotrons and possess similar characteristics. Other diodes contain both argon gas and mercury vapor. These tubes possess a low pick-up voltage (about 4 volts) and a low inverse peak (about 55 to 75 volts).

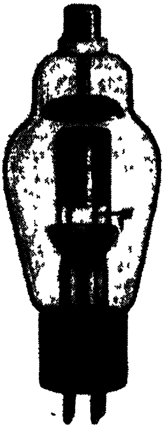
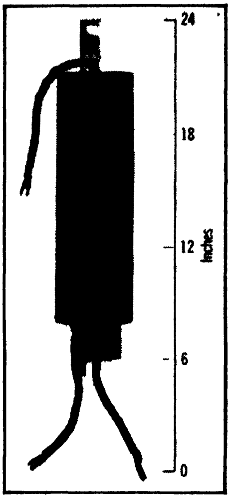


FIG. 6. Mercury-vapor rectifier tube (average anode current 0.25 ampere). (Courtesy Radio Corporation of America.)

Thyratrons. The thyratron is a triode containing inert gas or mercury vapor under low pressure. It is a phanotron plus a control grid. The admission of gas or vapor into a three-electrode vacuum tube greatly changes the characteristics and operation of the device. The presence of gas renders the tube useless as a grid-controlled amplifier but makes it valuable as a grid-controlled arc rectifier. In order to perform this new function the grid should surround the cathode or the anode so as to screen the one from the other. One early type of construction of the thyratron is shown schematically in Fig. 8.

The theory of action of the thyratron can be pictured by considering its behavior in a d-c circuit. Let the filamentary cathode of Fig. 8 be heated to normal temperature for emission and let switches S_p and S_c be open. Close S_p, placing a high positive potential on the anode with the grid free, and nothing will happen if the grid has a fine mesh. This follows because the free grid is made negative by the electrons from the cathode landing on it. This negative grid repels the electrons emitted by the cathode so that few get past the meshes of the grid to start ionization. Next, if switch S_c is connected to the positive side of the C battery, electrons pass through the grid readily, ionization starts, an arc develops, and the tube conducts an electric current which is limited by the magnitude of the load resistance. After the ionization has started, the opening of switch S_c, causing a free grid to exist, has no effect on the cathode-anode arc. Likewise, if the switch S_c is connected to the negative side of the C battery, the resulting negative charge on the grid has no appreciable effect on the arc. Thus the grid has the power to start the arc but no power to control its magnitude or to stop it after it has started. The reason for this unexpected action lies in

the presence of the positive ions in the arc. These ions, which fill the cathode-anode space, constitute the plasma of the arc. When the grid is made negative, it attracts some of the positive ions from the plasma



GENERAL CHARACTERISTICS

Number of Electrodes		2
Electrical		
Cathode-filamentary		
Filament voltage		2.5 volts
Filament current, approx		100 amp
Heating time, typical		2 min
Optimum phase of filament voltage with respect to anode voltage	.	90 degrees
Peak voltage drop, typical		9 volts
Rise of tube temperature above ambient		
Without forced-air circulation		
For average anode current	0	20 amp
Degrees rise, condensed mercury temperature	30	35
Degrees rise, temperature of side of tube	100	150
Maximum ratings		
Maximum peak inverse anode voltage		
20° to 60° C condensed mercury		1500 volts
20° to 70° C condensed mercury		800 volts
Maximum anode current		
Instantaneous		75 amp
Average		20 amp
Surge, for design only		750 amp
Maximum time of averaging current		30 sec
Temperature limits, condensed mercury +40° to +60° C		

FIG. 7. Phanotron FG-166. (Courtesy General Electric Company.)

and these positive ions fly to and cluster around the grid like bees coming to a hive (Fig. 9). Each positive ion seizes an electron and becomes a neutral atom. However, the process is continuous and the

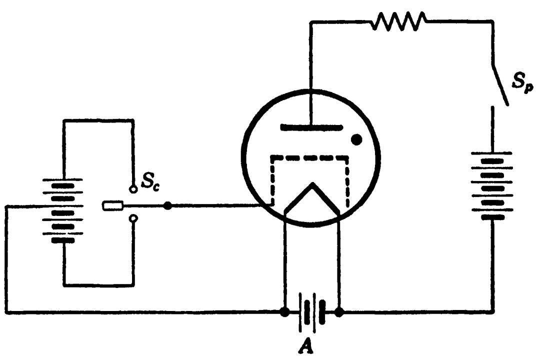


FIG. 8. Thyratron in a d-c circuit.

cluster or sheath of positive ions surrounding the grid creates a positive field (space charge) as far as the surrounding region is concerned. If the grid is made more negative, a thicker sheath of positive ions

surrounds it but without any effect on the cathode-anode arc, if the meshes of the grid are some distance apart. However, if the grid wires are very close together (almost touching), the sheath of positive ions around the grid will serve to limit the cathode-anode arc. Thus, theoretically, a closely spaced grid plus a high negative grid potential can be used to limit and even to stop an arc. Such means are not employed in practice. The simple and usual method for interrupting the arc when using direct current is to break the cathode-anode circuit by opening the switch S_p. The arc will be interrupted also if the plate voltage is reduced below the minimum required to maintain ionization.

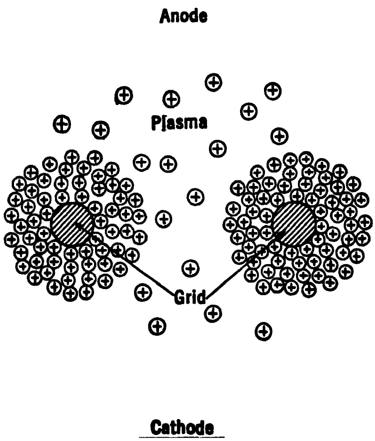


FIG. 9. Grid action in a thyratron.

The potential required on the grid of a thyratron to permit the rectifying arc to start depends on several factors. These factors are the size of the openings of the grid, the gas or vapor pressure within the tube, the general geometry of the tube, the grid current and circuit resistance, and the potential applied to the anode. The voltage conditions for starting the arc depend to a large degree on the structure of the grid. Thus with a fine-mesh grid (small hole or holes) a positive potential must be applied to the grid to start gaseous conduction, whereas with a coarse grid (large hole or holes) the arc may start with a negative potential on the grid. For the range of gas pressures used on thyratrons the low pressures require a higher potential gradient to start the arc. The positive potential on the anode determines the electric field within the tube and hence the "pull" upon the electrons. Thus a high initial positive voltage on the anode will require a more negative potential on the grid to prevent the formation of an arc. For a given tube, there is a certain ratio of anode volts to grid volts at which the tube will "fire." This ratio is called the grid-control ratio. It can be expressed as

$$= -\left(\frac{e_a}{e_c}\right) \qquad i_a = 0$$

and it bears a resemblance to the amplifying factor mu of the vacuum triode. The grid-control characteristics of thyratrons from which the grid-control ratio can be obtained are illustrated in Fig. 10. The grid-control ratio for a tube with the characteristic A is cy/dy, and this ratio will be constant except for the lower portion of the curve.

Thyratrons may be classified by the sign of the grid-control voltage that permits the starting of the tube. A negative-control tube is illustrated by curve A of Fig. 10 where a negative potential must be applied to the grid to prevent starting for all values of anode potential within the range of operation. For a similar reason, curve C represents a tube having a positive control. Curve B of Fig. 10 represents a tube of the intermediate-control class. This tube requires a positive grid for starting for anode voltages up to point X and a negative grid for higher values. This class of tube has a short deionization time and is used for inverter circuits.

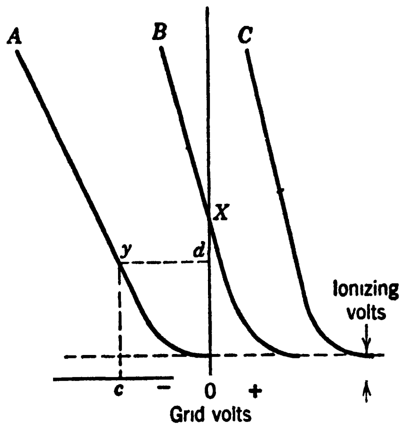


FIG. 10. Anode grid-voltage characteristic of thyratrons.

The thyratron is a "natural" for use on a-c circuits. If the grid is given a potential sufficiently positive so that the tube will conduct for all values of anode potential (above ionizing potential), the tube will rectify all positive loops of a-c potential exactly like a gaseous diode. However, if the potential of the grid of the thyratron is controlled in a suitable manner the thyratron may be made to conduct current or to stop current at will. This follows because the a-c voltage passes through zero twice during each cycle and is negative for half of each cycle. Thus if the grid of the thyratron is brought to the critical value, as determined by the grid-control ratio, the arc is struck each time the anode voltage becomes positive, but, if the grid voltage is lowered (made more negative), the current will stop at the end of the positive-anode voltage loop. Then when the anode voltage becomes positive on the next and succeeding cycles, the tube does not "fire." This simple grid control for starting and stopping current rectification

in the anode circuit has given rise to the term "trigger tube" for the thyratron. The operation of the thyratron on alternating current is illustrated in Fig. 11. The grid voltage necessary to stop conduction for any given anode potential is given by the grid-control characteristic on the left. The necessary restraining values can be referred to the impressed a-c voltage cycle as shown by the graphical construction of Fig. 11. The curve e_c shows the necessary values of grid po-

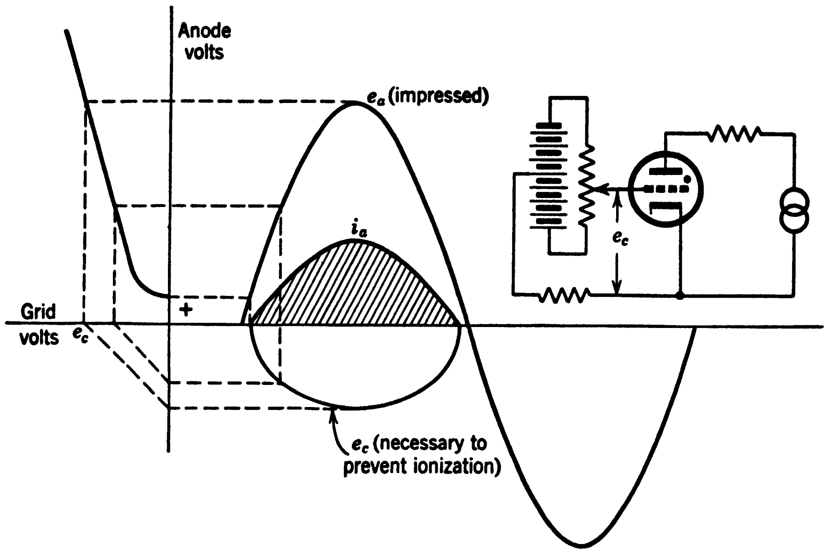


FIG. 11. Action of a thyratron on alternating current.

tential to prevent conduction for corresponding values of positive anode potentials. For a complete "stop" current, the grid must have a negative value greater than the maximum of the e_c curve.

The e_c curve of Fig. 11 suggests that the grid control of the thyratron may be used to control the magnitude of the rectified current as well as for "on" and "off" operation. Thus if the grid is held sufficiently negative until a certain part of each positive-anode voltage loop is reached, the tube will fire and conduction will occur for the remainder of each positive half of the cycle. The grid may be made to exercise this form of control in two ways, known as *amplitude control* and *phase-shift control.* Amplitude control is brought about by applying a negative bias to the grid as illustrated in Fig. 12. For the bias shown under *a* the thyratron conducts current beginning at point *X* and the area of the rectified current loop is reduced slightly. An increase in the

bias to the value shown under b will permit the tube to fire at point Y so that the magnitude of the rectified current has been reduced to one-half the normal value. Values of negative bias between those shown will give other magnitudes of current. Obviously, this is a very simple method of controlling current values but it has some limitations and disadvantages. First, the maximum reduction in magnitude is one-half value (one-fourth cycle). This maximum reduction is critical

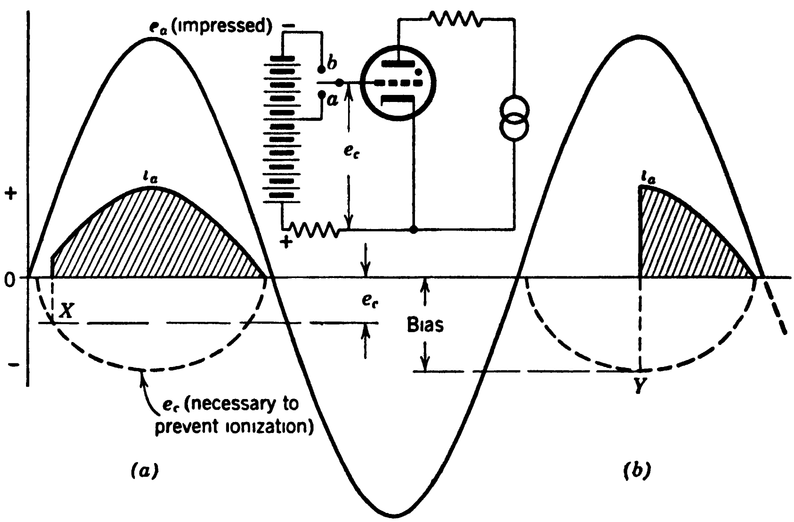


FIG. 12. Amplitude grid control of a thyratron.

because the firing point depends on both the grid and anode potentials and a temporary decrease in anode voltage might cause the tube to skip some cycles. For like reasons changes in anode potential, thermionic emission from cathode, and changes in temperature will vary the firing point, and hence the control of the current magnitude for a given grid-bias setting may not be sufficiently exact for some applications.

A more exact timing and firing of the thyratron can be secured by phase-shift control. In this method an a-c voltage derived from the anode supply line is applied to the grid. The phase of the grid voltage is shifted with respect to the anode voltage by suitable means, giving results as shown in Fig. 13. If the grid voltage e_g is "in phase" or ahead of the anode voltage as shown in part a, the tube fires as soon as the positive loop of the anode potential reaches the ionizing voltage and the complete loop is rectified. A delayed shift of the grid voltage e_g as shown in part b will cause the rectification to begin at the end of one-

third of the positive loops. A further shift in the grid voltage will delay the firing to a point near the end of the positive loop, as illustrated in part *c*. A shift of the grid potential from "in phase" to a lag of 180 degrees will permit a control from "full" to zero magnitude of current. Any change of anode voltage will be reflected in a change in grid

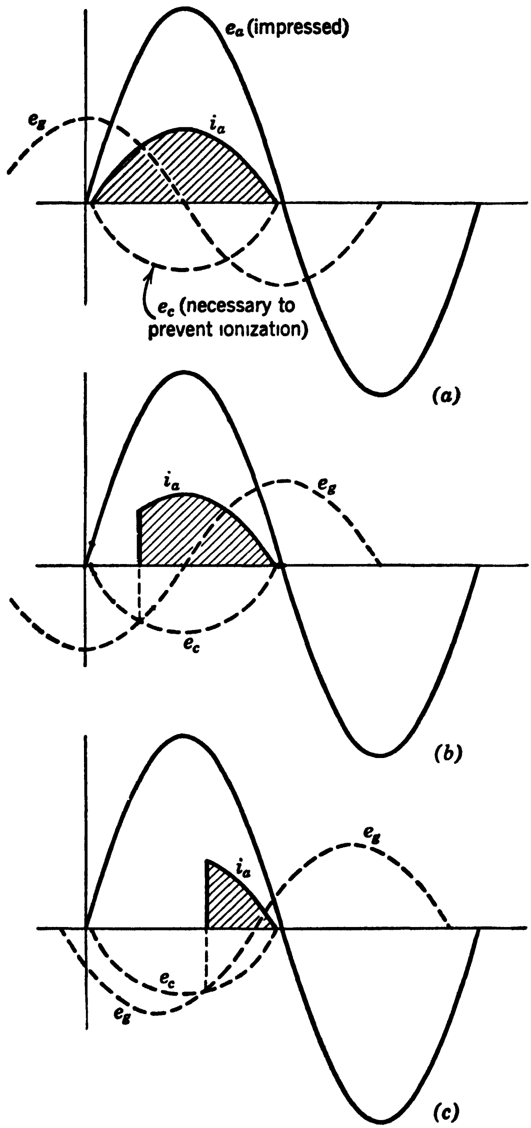


FIG. 13. Phase-shift grid control in a thyratron.

voltage and in the value of the e_c curve so that the period in the cycle where the firing occurs will remain nearly constant. For some applications it is advantageous to superimpose the controllable a-c shift component of the grid voltage upon a normal grid bias, as shown in Fig. 14. Here the grid bias is made sufficiently negative to prevent the tube from firing if the a-c grid potential is not applied. Since the grid is always negative, the power required for the grid input is reduced to a very low value. Another useful combination of a-c and d-c grid voltages for firing thyratrons is to use an a-c voltage component with a fixed phase shift (such as 90 degrees) and then control the firing angle by a variation of the d-c bias (vary the grid bias in Fig. 14).

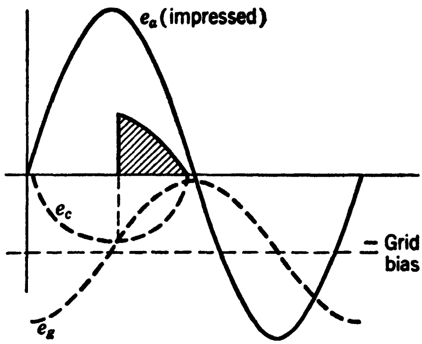


Fig. 14. Phase-shift control superimposed upon a normal grid bias in a thyratron.

The phase shift for the grid circuits of thyratrons and other electronic devices may be produced by mechanical devices or by simple circuits, one of which is illustrated in Fig. 15. The theory of this phase-shifting circuit is illustrated in the vector diagram of Fig. 15. The a-c voltage impressed across the cathode-anode circuit is in phase

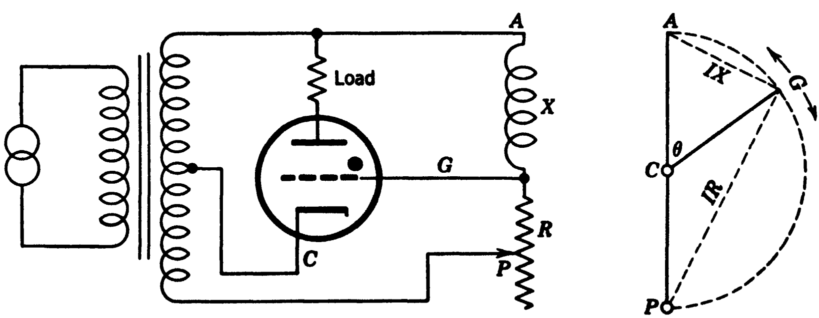


Fig. 15. Simple phase-shift circuit and vector diagram.

with that across points A and P. With a very high value of resistance R, the point G on the vector diagram will show voltage CG (cathode-to-grid) nearly in phase with that across the cathode-anode circuit.

Now as R is reduced, the magnitude and phase of AP remains unchanged but the position of G will swing clockwise on the arc of a circle, thus throwing the cathode-grid voltage out of phase with the cathode-anode voltage and making possible a 180-degree phase shift. For all positions of shift the magnitude of CG remains constant. A condenser may be substituted for the inductance of Fig. 15 to give another simple phase-shifting circuit. In making this change, the relative positions of R and C must be interchanged in order to retain a lagging phase shift.

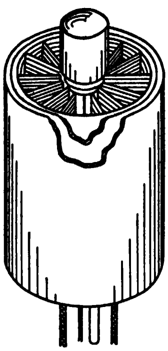


FIG. 16. Heater type of cathode used on thyratrons.

Some small-capacity thyratrons have cathodes of the filament type as in vacuum tubes. For the larger thyratrons there are two objections to this type of cathode. In the first place, the filament has a large heat loss and is very inefficient. In the second place, the voltage drop between the ends of filament reduces the available range of the cathode-anode potential drop. The neutralization of the negative space charge by the presence of positive ions permits a greater distance between cathode, grid, and anode and also makes possible the confinement of the cathode in a small space or in a heat-insulated oven. Several examples of cathodes used in thyratrons are illustrated in Fig. 5 and an additional shield and heater type is shown in Fig. 16.

The grids of thyratrons are placed farther from the cathode than in vacuum tubes to make the grid control more effective and to keep the grid at a lower temperature so that grid emission will not render the control ineffective. Several types of grids have been used in thyratrons. Grids in early tubes were hollow cylinders with one end closed. The walls of the cylinders consisted of (1) a wire mesh, or (2) sheet metals with small holes for the passage of electrons and ions (parts *b* and *c*, Fig. 17). The basic type for the grids used today consists of a hollow metal cylinder with open ends but closed by a baffle near the center. The anode and cathode are placed within the cylinder, the former above the baffle and the latter below. Holes for the passage of electrons and ions may be placed in the sides of the grid cylinder or in the baffle plate only (parts *c* and *d*, Fig. 17). For thyratrons in metal enclosures, the enclosure itself constitutes the grid structure.

Inert gas such as argon or mercury vapor is used in thyratrons at pressures varying from 1 to 50 microns (a micron is 10^{-6} meter of mercury). It is essential that the gas be pure. In normal operation

of a mercury-vapor arc there are 10^{+10} to 10^{+12} ions present per cubic centimeter. The gas in the tube aids in carrying the heat away from the anode.

Thyratrons are built in many sizes and ratings. Those of low capacity are contained in glass tubes and use inert gas. Those of large capacity are contained in metal tubes and use mercury vapor. It is possible to build thyratrons with current ratings as high as 100 amperes and with voltages up to 100,000. However, for large current capacities it has been found preferable to use other types of tubes to carry the load current and to utilize the thyratron for controlling the second tube. A miniature thyratron for control is illustrated in Fig. 18. The construction, rating, and characteristics of two typical larger thyratrons are given in Figs. 19 and 20.

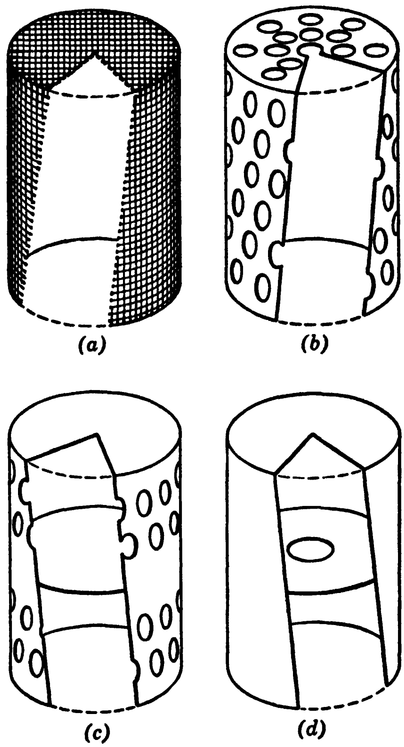


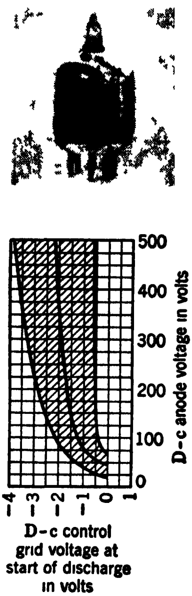
FIG. 17. Grid structures for thyratrons.

The efficiencies of thyratrons in voltage ratings of 500 and up are of the order of 97 per cent, which is excellent compared with other electrical machines and devices.

The precautions previously stated regarding the operation of mercury-vapor diodes should be observed in the use of thyratrons. The cathode should be brought up to normal operating temperature before voltage is applied to the anode to prevent cathode sputtering. The output current must not exceed the emission current. The cathode-anode voltage must be kept below the disintegrating potential for the cathode (about 28 volts). Both the current output and the cathode-anode voltage should be limited by a suitable load impedance.

It is usually necessary to place a resistor in the grid circuit of three-electrode thyratrons. There are a number of reasons for this practice. Since the grid lies within the path of the electron and ion flow, a fairly

large current may flow in the grid circuit for the tube that requires a positive grid or for any tube during the ionization period. When an a-c voltage is impressed across the grid, a grid current flows because of a rather large interelectrode capacity between the cathode and grid. Grid current in the thyratron produces some heating, changes the input impedance to the grid, and is likely to interfere with the operation



GENERAL CHARACTERISTICS

Number of electrodes	4
Electrical Design	
Cathode—indirectly heated type	
Voltage	6.3 volts
Current, approx	0.15 amp
Heating time, typical	10 sec
Peak voltage drop, typical	11 volts
Average anode-to-control-grid capacitance	0.1 $\mu\mu$f
Mechanical Design	
Net weight, approx	⅛ oz
Operating position—any	
Maximum overall length	1⅜ in
Maximum overall diameter	1$\frac{1}{16}$ in
Maximum ratings	
Maximum peak anode voltage	
Inverse	500 volts
Forward	500 volts
Maximum anode current	
Instantaneous	100 ma
Average	20 ma
Maximum time of averaging anode current	15 sec
Ambient temperature limits	$-40°$ to $+80°$ C

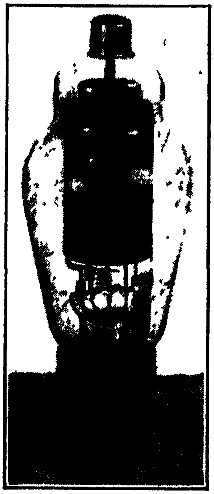
FIG. 18. Miniature thyratron GL-546. (Courtesy General Electric Company)

of the grid-control circuit. Grid series resistors have magnitudes in the range of ¼ to 2 megohms.

The thyratron with three electrodes has certain limitations somewhat similar to those of the vacuum triode which can be overcome through the use of a fourth electrode called a shield grid corresponding to the screen grid in the vacuum tetrode. The shield grid is usually a hollow metal cylinder containing two circular disk baffles as shown in Fig. 21. Each baffle contains a hole for the passage of electrons and ions. The control grid is a small ring placed in the conduction path midway between the baffles of the shield grid. The shield grid is maintained at a constant d-c voltage which tends to stabilize operation and to give the desired grid-control ratio.

The shield-grid construction serves to minimize grid current from a variety of causes and to permit satisfactory operation of the tube in a

high-impedance grid circuit. These advantages are accomplished in three ways. First the shield grid reduces both the anode-to-control-grid and the control-grid-to-cathode interelectrode capacity. The former is important in thyratron circuits since the steep wave-front



GENERAL CHARACTERISTICS

Number of Electrodes			3
Electrical Design			
Cathode—filamentary			
Filament voltage			2.5 volts
Filament current, approx			5.0 amp
Heating time, typical			5 sec
Peak voltage drop, typical			16 volts
Approximate control characteristics			
Anode voltage	40	100	1000 volts
Control-grid voltage	0	−2.25	−6.5 volts
Anode-to-control-grid capacitance			4.4 $\mu\mu$f
Deionization time, approx			1000 μsec
Ionization time, approx			10 μsec
Maximum ratings			
Maximum peak anode voltage			
Inverse			5000 volts
Forward			2500 volts
Maximum negative grid voltage			
Before conduction			500 volts
During conduction			10 volts
Maximum anode current			
Instantaneous, 25 cycles and above			2.0 amp
Instantaneous, below 25 cycles			1.0 amp
Average			0.5 amp
Surge, for design only			40 amp
Duration of surge current			0.1 sec
Maximum grid current			
Instantaneous			0.25 amp
Average			0.05 amp
Maximum time of averaging current			15 sec
Temperature limits, ambient			+40° to +80° C

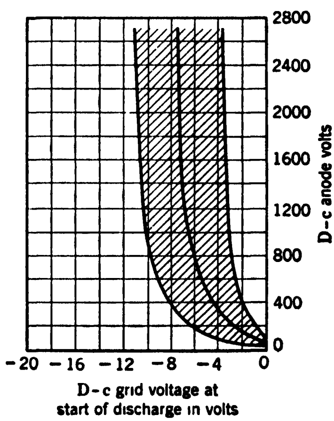
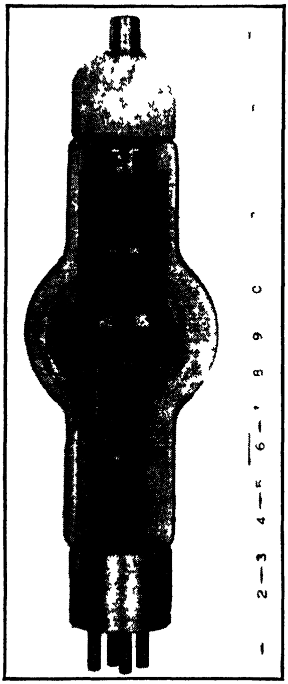


FIG. 19. Thyratron FG-17. (Courtesy General Electric Company.)

transients frequently encountered in these circuits may be transmitted to the control grid through this capacity causing premature firing. The reduction in the cathode-to-control-grid capacity is of lesser importance though it does reduce the grid current input where a-c potentials are applied. The second advantage of the shield grid is its action in

shielding the control grid from contamination and temperature. The shielded position reduces the amount of material evaporated or sputtered from the cathode and anode which may become deposited on the control grid. Also the shielded position reduces the radiated heat



GENERAL CHARACTERISTICS

Number of Electrodes				3
Electrical Design				
Cathode—indirectly heated type				
Heater voltage				5.0 volts
Heater current, approx				20 amp
Heating time, typical				5 min
Tube voltage drop, typical				16 volts
Approximate starting characteristics				
Anode voltage	1,000	10,000	250 approx	volts
Control-grid voltage	−1.5	−4.5	0	volts
Deionization time, approx				100 μsec
Ionization time, approx				20 μsec
Maximum Ratings				
Maximum peak anode voltage				
Inverse				10,000 volts
Forward				10,000 volts
Maximum negative control-grid voltage				
Before conduction				1,000 volts
During conduction				15 volts
Maximum anode current				
Instantaneous, 25 cycles and above				75 amp
Instantaneous, below 25 cycles				25 amp
Average				12.5 amp
Maximum control-grid current				
Instantaneous				5.0 amp
Average				1.0 amp
Maximum time of averaging current				30 sec
Temperature limits, condensed mercury				40° to 65° C

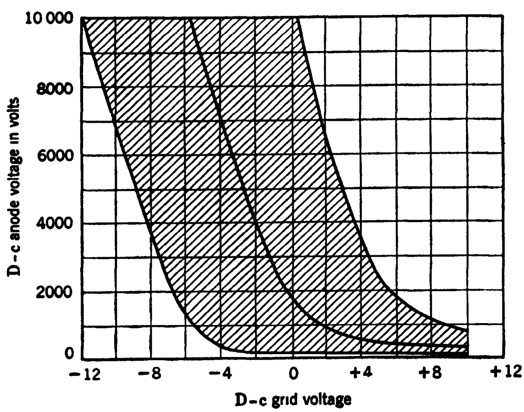


FIG. 20. Thyratron FG-41. (Courtesy General Electric Company.)

and lowers the temperature of the control grid which, in turn, reduces grid emission. A third advantage of the shield grid is to permit the use of a small control grid which reduces both the emission current and any current arising from the interelectrode capacity. It should be noted that the shield-grid current does not pass through the control-grid circuit.

GENERAL CHARACTERISTICS

Electrical Design		
Number of electrodes		4
Cathode type—indirectly heated		
Voltage	5.5	5.0 volts
Current, approx	11.0	10.0 amp
Heating time, typical		5 min
Peak voltage drop, typical		16 volts
Approximate starting characteristics		
Anode voltage	100	2000 volts
Shield-grid voltage	0	0 volts
Control-grid voltage	+1.0	−14.0 volts
Anode to control-grid capacitance		0.07 μμf
Deionization time, approx		1000 μsec
Ionization time, approx		10 μsec
Maximum Ratings		
Maximum peak anode voltage		
Inverse	750	2000 volts
Forward	750	2000 volts
Maximum negative control-grid voltage		
Before conduction		1000 volts
During conduction		10 volts
Maximum negative shield-grid voltage		
Before conduction		300 volts
During conduction		5 volts
Maximum anode current		
Instantaneous, 25 cycles and above	77	40 amp
Instantaneous, below 25 cycles		13.0 amp
Average	2.5	6.4 amp
Surge, for design only		400 amp
Duration of surge current		0.1 sec
Maximum control-grid current		
Instantaneous		1.0 amp
Average		0.25 amp
Maximum shield-grid current		
Instantaneous		2.0 amp
Average		0.50 amp
Maximum time of averaging current		15 sec
Temperature limits, condensed mercury	+30° to +95°	+40° to +80° C
Recommended temperature, condensed mercury		+40° C

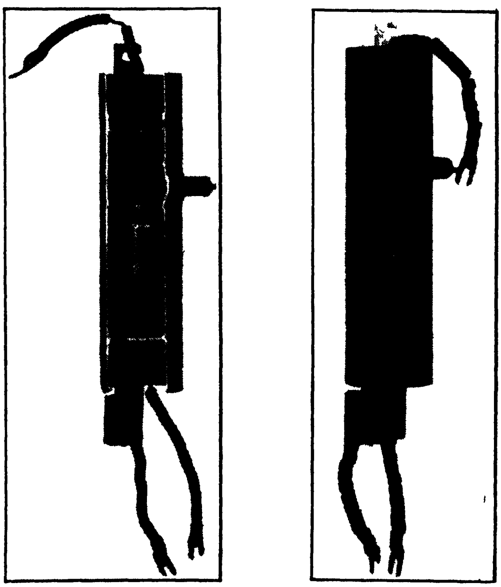


FIG. 21. Shield-grid thyratron FG-172. (Courtesy General Electric Company.)

The thyratron is one of the most useful control devices invented in the twentieth century. Its applications are too numerous to mention. In one application it serves as a relay. A small change of potential on the grid starts or stops a rectified current which is performing some useful function. A constant temperature is maintained in an electric furnace or oven by a thermostat which controls the phase shift or the on-and-off potential on the grid of the thyratron which furnishes rectified current to the device. Lighting circuits may be dimmed or lighted slowly or turned on and off by thyratrons. Here the thyratron varies the flux (saturation) in an iron-core reactor which is in series with the lighting circuit. The armature of a d-c motor may be supplied with direct current rectified by a thyratron, and the starting, stopping, and speed control may be governed by the voltage phase shift on the grid of the tube. The grid circuits of thyratrons may be energized by light falling upon a photocell.

An important factor in the use of the thyratron as a trigger tube is the time required for ionization and deionization. Ionization time for commercial thyratrons varies from 10 to 20 microseconds and the deionization time from 100 to 1000 microseconds. While these periods appear to be short, they are long enough to limit the use of the thyratron for many high-frequency applications. It is possible to reduce the ionization time for mercury-vapor thyratrons by using peaked grid voltages of relatively large magnitude for firing the tube. Where very short ionization and deionization time is important the hydrogen thyratron may be used.

The hydrogen thyratron is a hot-cathode grid-controlled gas rectifier tube developed during World War II for pulsing service at high repetition frequencies, high peak currents, and high voltages. Its outstanding feature is the short deionization time required to convert the gaseous ions (hydrogen) to neutral molecules when the tube is shut off. This action results from the relatively small mass of the hydrogen molecule. Another advantage of the tube is that it may be operated over a wide range of ambient temperatures without significant change in electrical characteristics. The hydrogen thyratron has been applied in high-frequency and pulsing service but may find other applications where its characteristics are useful. A commercial hydrogen thyratron and its rating are shown in Fig. 22.

Mercury-Pool Tubes. A mercury-pool tube uses a pool of cold mercury for a cathode. The conducting medium consists of emitted electrons plus ions produced in the mercury vapor. The conduction due to the mercury vapor is of the arc-discharge type. Since electrons

are not readily extracted from a cold metal, some special means must be employed (1) to initiate and (2) to maintain an emission of electrons from the mercury pool. Two general methods are employed for initiating the electron emission, one mechanical and one electrical. In the application of mechanical methods an auxiliary anode is caused

Peak anode voltage, max	8.0 kv
Peak anode current, max	90 amp
Peak inverse anode voltage, max	6.0 kv
Average anode current, max	100 ma d-c
Pulse duration (measured at ½ amplitude), max	6.0 μsec
Pulse repetition frequency, max	4000 pps

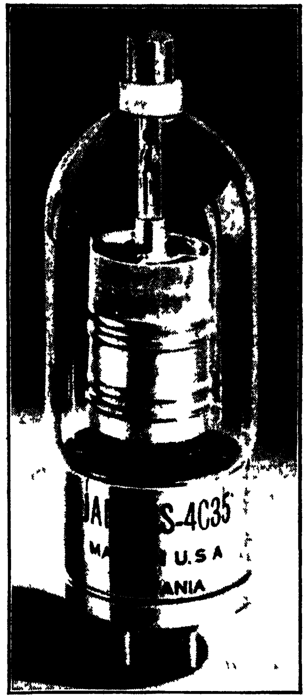


FIG. 22. Hydrogen thyratron. (Courtesy Sylvania Electric Products, Inc.)

to touch the mercury pool for an instant. This may be accomplished (1) by tipping the tube so that the liquid mercury flows to a point where cathode and anode are brought in contact, (2) by moving the anode so as to dip into the mercury pool, or (3) by squirting the liquid mercury upward so that it contacts an auxiliary anode momentarily. In any of these processes the momentary contact between the mercury (cathode) and anode causes a transient current followed by a spark and perhaps a small arc which results in an initial emission of electrons from the mercury. One electrical method for starting electron emission is to place an auxiliary anode (called an ignitor) so that its point dips into the mercury. Then a transient current from point to mercury

produces a spark for starting emission. In some mercury-arc lamps a very high transient electric field has been created close to the mercury pool by an inductive voltage "kick." This transient field has been sufficient to cause a breakdown and start of emission on the glow-discharge principle. After the emission of electrons is initiated the maintenance of emission depends on additional theoretical considerations. As soon as electron emission is initiated, the electrons are attracted toward the main anode and produce ionization. An arc develops and the plasma of positive ions fills most of the space between the anode and the mercury pool. The current seems to originate at one or more hot spots on the mercury pool. The hot spots look like little craters on the surface of the pool and they travel about over various paths or eddies. The craters are formed by the bombardment of the positive ions on the mercury. At first it was assumed that the temperature of the hot spots was sufficient to produce thermionic emission from the mercury. This concept has been abandoned because the temperature necessary for thermal emission would vaporize all the mercury and that does not happen. Accordingly, the theory that has rather general acceptance is that the positive ions form a very high positive space charge just above the mercury. The resulting sheath between the mercury and plasma is so thin that the electric field in this sheath is sufficient to secure emission through the principle of high-field emission.

One important advantage of all mercury-pool tubes is that the mercury cathode is capable of furnishing enormous emission temporarily without damage. Thus the tube will withstand temporary overload and even short circuit without destruction of the cathode. The mercury vapor condenses on the walls of the tube and returns to the pool.

One type of mercury-pool rectifier has been used since the early part of this century. This device and its circuit are shown in Fig. 23. The tube was built with two anodes for full-wave rectification. The glass enclosure had a large upper chamber to provide cooling area. The arc was started by tilting the tube so that the mercury could contact the auxiliary starting anode. These tubes had a maximum capacity of about 30 amperes and were widely used for charging storage batteries. A few tubes of this type are still in service for rectifying alternating current for use on series d-c street-lighting systems of the luminous-arc type. These early pool-type rectifiers had four disadvantages: (1) They had a fragile glass envelope; (2) they were not suited to the dissipation of heat; (3) they were limited in current capacity; and (4) their rectified voltage and current waves had wide fluctuations and were not suitable where a smooth flow of d-c power was

desired. These disadvantages led to the development of the water-cooled metal-tank multielectrode rectifier.

Multielectrode Metal-Tank Rectifier. The multielectrode rectifier was introduced into the electric-power field about 1925. This rectifier consists of a large cylindrical steel tank surrounded by a water-cooled jacket (Fig. 24) and evacuated to a pressure of one-millionth of an atmosphere. The cathode consists of a pool of mercury at the bottom (Fig. 25). There are from six to eighteen main anodes, like the one shown at the left of the figure. It consists of a graphite cylinder connected to the outside through an insulating bushing in the wall of the tank. The main anodes are placed in a cylindrical insulating shield. The value of using multielectrodes is twofold. First, electricity is transmitted most economically by multiphase circuits (usually three-phase), and, secondly, the multiphase rectifier gives a smoother output. The latter statement is illustrated by the rectified wave forms of Fig. 25. The single-wave, three-phase circuit gives a more desirable output wave, and the additions of more phases and anodes will increase the number of ripples but reduce their magnitude. A three-phase supply may be split into six, twelve, and eighteen phases by suitable transformer connections. The steel-tank, mercury-arc rectifier is a very rugged device, effectively cooled by water for any desired capacity and designed to give a smooth voltage, current, and power output.

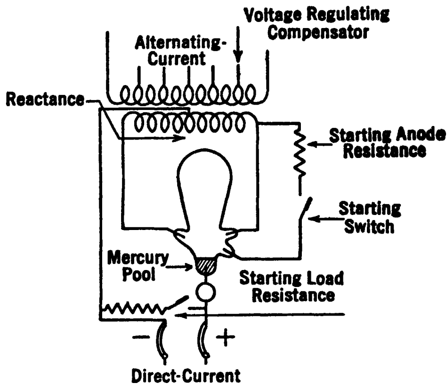


FIG. 23. Full-wave mercury-arc rectifier.

The steel-tank rectifier is started by the small central starting anode (Fig. 25)which can be lowered into the mercury and then withdrawn by an electromagnet. The breaking of the mercury contact on withdrawal strikes the arc which is then "picked up" by one of the main or auxiliary anodes which has the highest positive potential at that instant. The various main anodes will pick up the arc in rotation as they become sufficiently positive and maintain it until the load is taken over by another of higher positive potential. The output current at any instant is the arithmetic sum of the current of all anodes. Theoretically, only one anode having the highest positive potential would

be expected to carry the load current at a given instant. In practice the inductance in transformer windings in series with the anodes causes an anode to carry current for a time after a succeeding anode begins to conduct. Hence in practice, two or more of the multianodes

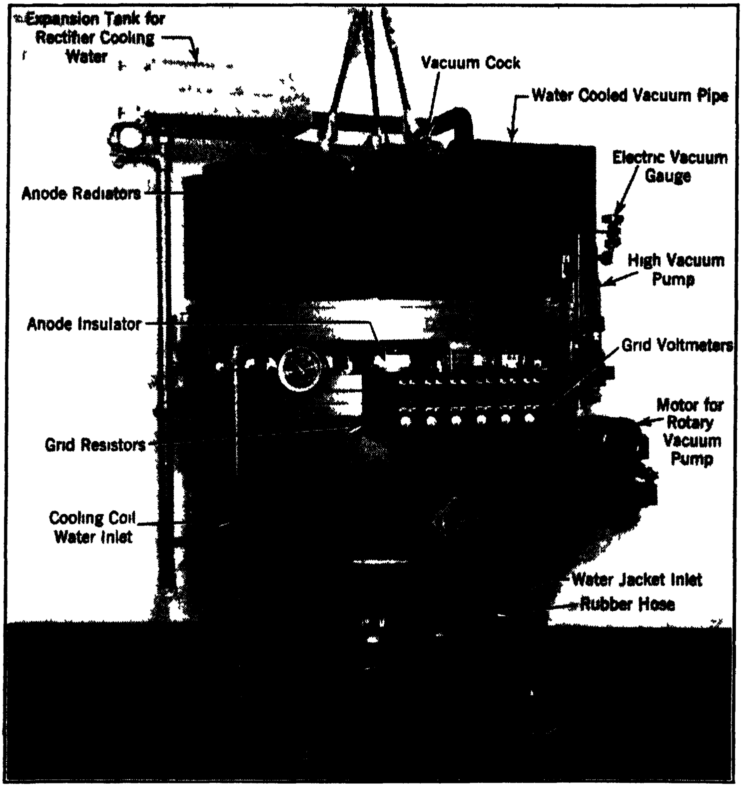


FIG. 24. Steel-tank mercury-arc rectifier. (Courtesy Allis-Chalmers Manufacturing Company.)

may be carrying part of the load current simultaneously. In the use of the multielectrode rectifier, the rectified d-c load may go to zero, which would extinguish the arc and make it necessary to re-establish ionization whenever the load returns. To avoid this contingency, one or more auxiliary anodes (see right side of Fig. 25) are constantly energized through a circuit independent of the load and they serve to keep gas ions present in the tank continuously. The continuous presence of gas ions in the arc chamber introduces a problem in the operation of the multielectrode rectifier. When any one of the multianodes

has a negative inverse voltage on it, it repels electrons and, theoretically, its current goes to zero, but positive ions may be present near it due to electrons en route to an adjacent anode. These positive ions will be attracted to the first anode and their bombardment may pro-

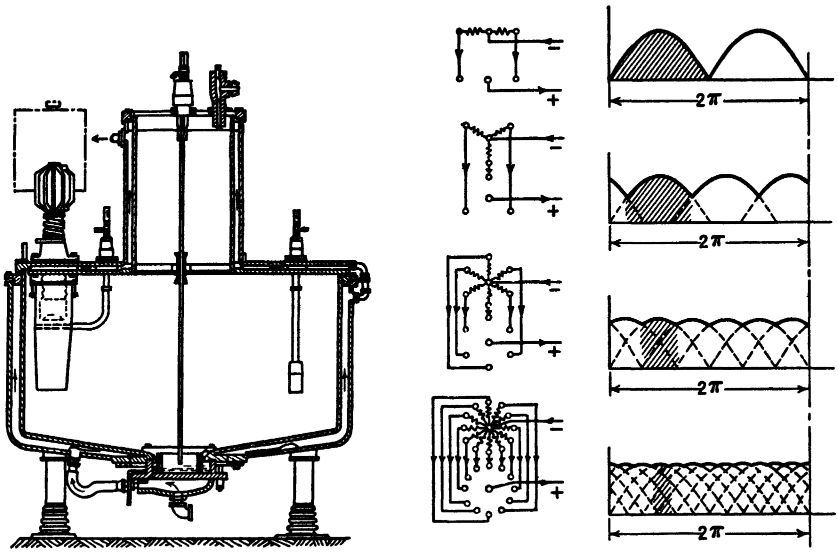


FIG. 25. Cross section of a steel-tank mercury-arc rectifier and rectified wave forms from simple and multiphase rectifiers. (Courtesy Allis-Chalmers Manufacturing Company.)

duce a hot spot on the anode which will emit electrons. Then this anode becomes a cathode and a reverse arc starts which will become a short circuit from anode to anode in the rectifier. Such a short circuit is called an "arc-back" or a "flash-back" and it will open the protective equipment and cause a shutdown.

The tendency toward arc-back is greatly reduced by the presence of a cylindrical shield placed around each anode. When the anode is positive, positive ions may be formed within the shield. However, as the potential falls to zero the positive ions are drawn out of the shield toward the cathode pool at the bottom of the tank or deionized by the shield. Hence the space within the anode shield becomes deionized and there is little tendency for arc-back. Occasional arc-backs do occur in the multielectrode mercury-arc rectifier.

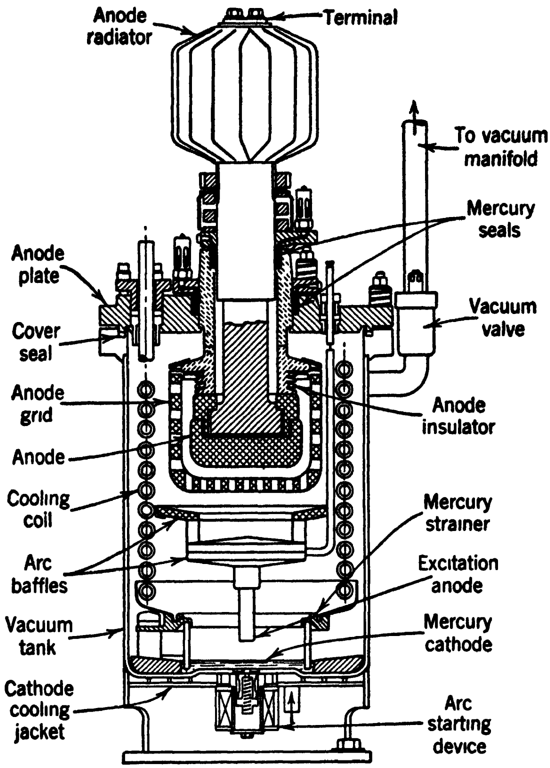
The principle of grid control as used in other electronic devices may likewise be employed in the multianode rectifier. Here the grid is

placed within the grid shield chamber and close to the anode (see Fig. 25). In this position the grid can control the time of the starting of electron flow to the anode on each positive voltage wave. A negative voltage on the grid prevents starting, and the time when the grid potential changes to zero or positive may be controlled through a phase shift of the voltages supplied to the rectifier. This grid action controls the rectified output voltage and may be used to control output current for special applications.

The multielectrode steel-tank rectifier is widely used for large-power d-c applications. Over 3,000,000 kilovolt-amperes of these rectifiers are in service. The important applications are electric railway service, electrolytic processes in industry, and motive power in steel mills. The advantages of these units over rotary types of conversion are (1) simple and rugged construction, (2) long life, (3) high overall conversion efficiency, (4) high momentary overload capacity, and (5) quiet operation. Certain disadvantages of the multielectrode tank rectifier are causing it to be superseded by the single-anode type described on the following pages. These disadvantages are: (1) There is a greater possibility of arc-backs due to several anodes in the same arc chamber. (2) The multianodes require large spacing between cathodes and anodes with correspondingly higher arc drops and reduced efficiency. (3) The damage to any anode or part causes the entire unit to be taken out of service for repairs.

Excitron Single-Anode Rectifier. The excitron is a single-anode rectifier embodying the same general principles of construction and operation as outlined for the multianode unit. A cross section of an excitron unit is shown at the top of Fig. 26. The unit is placed in operation by a small jet of mercury thrown upward upon the excitation anode by the arc-starting device. As the mercury falls away a pilot arc is initiated and maintained by the d-c supply from the selenium rectifier and control circuit illustrated in the lower part of Fig. 26. The anode is surrounded by a grid consisting of a perforated graphite cylinder which serves to time the period of current rectification. Grid action also combines with the restraining influence of baffles to prevent arc-backs due to inverse voltages. Arc-back from one main anode to another (as in multianode rectifiers) cannot occur.

Excitrons may be used in groups of three, six, twelve, or eighteen as required to serve the same as the multianode units. The individual units are rather light in weight and are easily installed and repaired. Since the anodes are closer to the cathodes the arc drop is



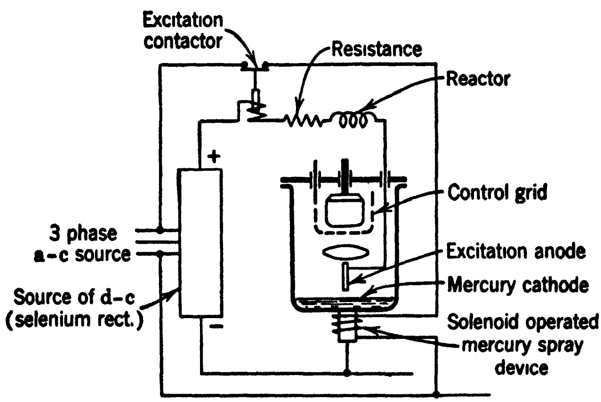


FIG. 26. Cross section and circuit of an excitron. (Courtesy Allis-Chalmers Manufacturing Company.)

lower and the efficiency is from 3 per cent to 4 per cent higher on output potentials of 250 to 300 volts.

Ignitron. The ignitron is a type of three-electrode mercury-vapor tube developed by Slepian and Ludwig in 1932. The name ignitron was derived from the novel method of igniting or starting the arc in this mercury-vapor tube. The ignitron has a graphite anode in the form of a disk or cylinder. Its cathode consists of a pool of mercury placed at the bottom of the tube. The third electrode, called the *ignitor*, terminates in a point which dips into the pool of mercury (Fig. 27). The ignitor may be a rod of ceramic material such as silicon carbide or one having a graphite center with a boron carbide coating.

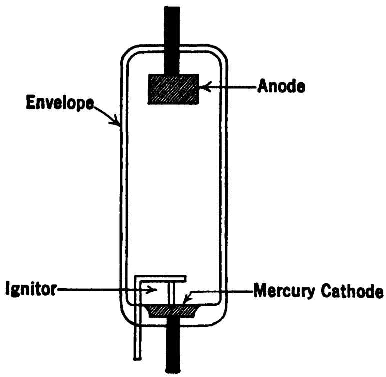


FIG. 27. An ignitron.

The theory of operation of the ignitron is simple. With a voltage placed across the cathode-anode circuit (anode positive) nothing happens because there are no electrons emitted from the cold mercury cathode. However, if a flash of current is passed through a circuit connecting the ignitor and cathode a spark will be created at their contact. This spark will produce some emission, and ionization of the mercury vapor will result. The arc thus established will continue as long as a suitable potential is maintained in the anode-cathode path.

The phenomena accompanying the development of the arc in the ignitron consist of two parts. The mercury does not come into intimate contact with the rough surface of the ignitor which consists (microscopically) of a number of sharp points. The flash of voltage across the ignitor-cathode circuit causes a potential gradient of approximately one million volts per centimeter across the point contacts, which is high enough to pull electrons out of the cold mercury. Simultaneous with this initiation of electron emission, the rise of current through the high resistance rod of the ignitor produces a potential gradient between the surface of the mercury and the top of the rod. This electric field accelerates the emitted electrons upward to ionization, and the development of an arc to the ignitor terminal which, in turn, is transferred upward to the anode (if positive).

The ignitron finds its application as a controlled rectifier for alter-

nating currents. Since the rectified anode current will go to zero each time the negative voltage loop is applied, it is necessary to ignite the tube for each positive cycle. This result can be obtained by the circuit given in Fig. 28. The hot-cathode diode in the auxiliary circuit conducts current as soon as its anode reaches the ionizing potential. This current passes through the ignitor, causing it to fire the ignitron since its anode becomes positive at the same time. When the ignitron fires, its arc-drop or cathode-anode potential falls to a low value of about 12 volts. This lowers the voltage across the series auxiliary

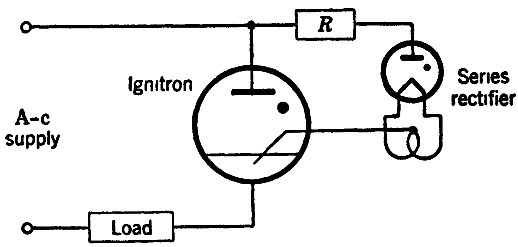


FIG. 28. Anode firing circuit for an ignitron.

circuit, consisting of rectifier and resistance at ignitor, so that the current in this circuit becomes very low or falls to zero.

The output of the ignitron may be regulated by controlling the time phase of firing or igniting. One method for accomplishing this control substitutes a thyratron for the diode rectifier in Fig. 28. The desired time of firing the ignitron is secured by either amplitude control or phase-shift control of the thyratron. Another method of firing and controlling the output of the ignitron utilizes special electrical networks having "trigger" action.

Ignitrons are often fired by the discharge of a condenser using some form of trigger circuit for initiating the discharge. One reactor-excitation circuit of this type is shown in Fig. 29. This circuit consists of three main parts: The firing circuit proper, a voltage-compensating network, and a phase-shifting reactor. In the firing circuit a capacitor C is charged through a linear reactor (constant reactance throughout range of applied voltage). The capacitor voltage is impressed across the ignitor circuit through a saturable reactor and directional filters (copper oxide rectifiers, page 185). The saturable reactor becomes magnetically saturated when the impressed voltage (and resulting current) reaches a predetermined value. At this critical point the reactance of the saturable reactor decreases and the capacitor discharges

a peak current through the ignitor-cathode circuit along the path shown by the arrows.

The function of the voltage-compensating network (C_1 and L_1 in parallel) is to furnish a nearly constant input voltage to the firing circuit regardless of fluctuation in the a-c supply. L_1 is a saturating reactor so designed that at normal line voltage its lagging circuit is just balanced by the leading current of capacitor C_1. If the line voltage is too high, the reactor L_1 begins to saturate and draws an excess mag-

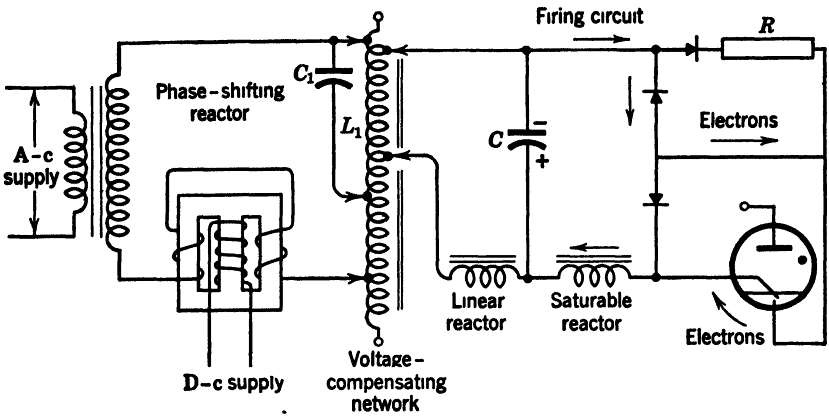


FIG. 29. Network for firing ignitrons (bold-face arrows show direction of conventional current flow).

netizing current. This lagging current passing through the phase-shifting reactor (left) produces a quadrature voltage drop and reduces the firing-circuit voltage. If the line voltage is low, L_1 draws a small magnetizing current and the stronger leading capacitor-charging current in passing through the phase-shifting reactor increases the firing-circuit voltage. This voltage-compensating action is so effective that the supply-line voltage can vary 50 per cent without ignition failure.

The phase-shifting transformer (often called a saturable reactor) has a three-legged iron core with series additive windings on the outer legs. With the d-c coil unexcited the series coils are highly inductive and will shift the phase of the voltage applied to the firing circuit. If sufficient direct current is passed through the coil on the central leg to saturate the entire core, the reactance of series windings falls to a low value and the phase of the voltage applied to the firing circuit is advanced. Intermediate values of direct current will give a corre-

sponding degree of phase shift which will determine the time of firing the ignitron.

The a-c supply for this reactor excitation must be the same as that applied to the anode-cathode circuit of the ignitron. The circuit of Fig. 29 will control single-phase, half-wave rectification. Full-wave rectification can be attained by substituting the ignitor-cathode circuit of a second ignitron for resistor R.

GENERAL CHARACTERISTICS

Mercury-pool triode	Clamp cooled
Line voltage range	250–600 volts
Tube voltage drop	12–18 volts
Ignitor voltage, max positive peak required	200 volts
Ignitor current, max peak required	30 amp
Ignition time, max	100 μsec
Net weight	1.5 lb
Shipping weight	3 lb

Maximum Anode Rating

Demand at 12.1 average amp *	300 kva
Demand at 22.4 average amp *	100 kva
Time of averaging current at 250 volts	22 sec
Time of averaging current at 500 volts	11 sec
Current surge peak at 250 volts	3360 amp
Current surge peak at 500 volts	1680 amp

Maximum Ignitor Rating

Voltage, max positive	900 volts
Voltage, max negative	5 volts
Current, max	100 amp
Current, 5-sec average	1 amp
Max temperature of cooling clamp	50° C

* Demand current and kva are on the basis of full-cycle conduction (no phase delay) whether or not phase control is used.

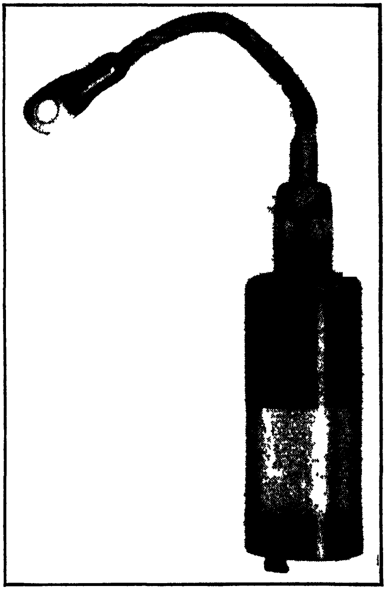
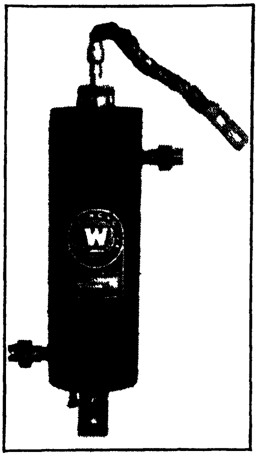


FIG. 30. Light-duty ignitron. (Courtesy Westinghouse Electric Corporation.)

The first ignitrons were encased in glass tubes as shown in Fig. 27. Later practice has tended toward building ignitrons in metal envelopes in order (1) to facilitate the dissipation of heat, (2) to increase their physical sturdiness, and (3) to improve the psychological reaction of industry toward the use of these tubes. The metal-encased ignitrons may be classified as to type of seal and type of cooling. In the smaller size the tube is evacuated and given a permanent seal (see Fig. 30). This type may rely on radiation and ordinary convectional circulation for cooling. The use of forced air circulation and also the addition of cooling fins will increase the rated capacity of these tubes. In the intermediate and large sizes the metal tubes are designed for cooling by water circulation through copper tubes or helical paths around the metal case (Fig. 31). In the large sizes the evacuation of the ignitron

is maintained by a mercury condensation pump. Although these larger tubes could be built with a permanent seal and thus avoid the cost and inconvenience of the evacuation equipment, another factor in economy demands the use of the nonseal type. This factor is the



IGNITRON WL-679

FOR RECTIFIER SERVICE

General Characteristics

Mercury-pool electronic tube—water cooled	
Tube voltage drop	12–20 volts
Ignitor voltage, max positive peak required	150 volts
Ignitor current, max peak required	40 amp
Ignition time, max	100 μsec
Temp rise, 150 amp average and water flow 1 5 gpm	7° C
Mounting position—vertical, cathode down	
Net weight	14 lb
Shipping weight	20 lb

Maximum Ratings 25 to 60 Cycles

Anode voltage, peak forward and inverse	900	2100 volts
Anode current, peak	900	600 amp
Anode current, continuous service, max average	100	75 amp
Anode current, 2-hour service, any 2-min average	150	112 5 amp
Anode current, 1-min service, any 1-min average	200	150 amp
Anode current, surge, 0 10-sec max	6000	4500 amp
Auxiliary anode voltage, peak inverse, main anode conducting	25	25 volts
Auxiliary anode voltage, peak inverse, main anode not conducting	150	150 volts
Auxiliary anode current, max average	5	5 amp
Ignitor voltage, max positive peak allowed	900	2100 volts
Ignitor voltage, max negative peak allowed	5	5 volts
Ignitor current, max peak allowed	100	100 amp
Ignitor current, max average 10-sec max averaging time	2	2 amp
Cooling water flow, min	1 5	1 5 gpm
Cooling water temperature max at outlet	60° C	45° C
Cooling water temperature min at inlet	10° C	10° C
Cooling water temperature range, optimum	20–45° C	20–40° C

For Welding Service

Line voltage, rms	2400 volts
Maximum demand at 75 average anode amp	1200 kva
Maximum demand at 113 average anode amp	600 kva

FIG. 31. Heavy-duty ignitron (Courtesy Westinghouse Electric Corporation)

uncertainty covering the life of the sealed unit and a preference for the opportunity to take the unit apart for repairs rather than to discard a very large and expensive device. An ignitron of large capacity is illustrated in Figs. 32 and 33.

It is logical to assume that the ignitron should withstand an infinite inverse voltage without arc-back. It would seem that all positive and negative ions would be swept out of the cathode-anode arc at the con-

clusion of the positive-voltage loop, and hence no inverse voltage could cause a breakdown. This logical theory was assumed by the early investigators of this device, but it has not been proved by practice. Occasional arc-backs do occur in ignitrons though they are not damaging to equipment. These arc-backs may be due to tiny hot spots which develop on the graphite anode and serve temporarily as emitters.

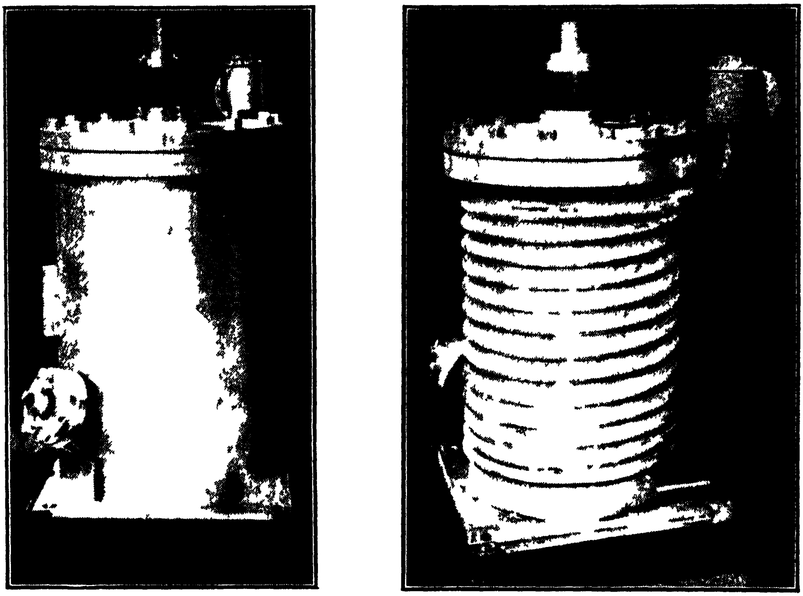


FIG 32 Large-size water-cooled ignitron (with and without outer case) (Courtesy Westinghouse Electric Corporation)

In order to improve the arc-back characteristic, baffles may be placed in the ignitron as illustrated in Fig. 33. By increasing the length of the arc path these baffles reduce the possibility of a reverse flow of current. An anode shield or shield grid placed around the anode with suitably timed potential will further serve to reduce any tendency to arc-back (Fig. 33).

There are two important applications of the ignitron in industry. The first is in the control of resistance welding—spot, butt, and line welding. The second application is for rectification for d-c power applications. The use of the ignitron has been rather revolutionary in making possible new applications of welding because it assures a uniformity of results never attained in the past. In this field a combination of two ignitrons serves as an electronic switch for controlling the

time of application of an alternating current in making resistance welds. A simplified circuit for performing this function is given in Fig. 34. Two ignitrons are connected with reversed polarity in parallel with each other and in series with the primary of a welding transformer. The two ignitors are connected in series with each other and with a

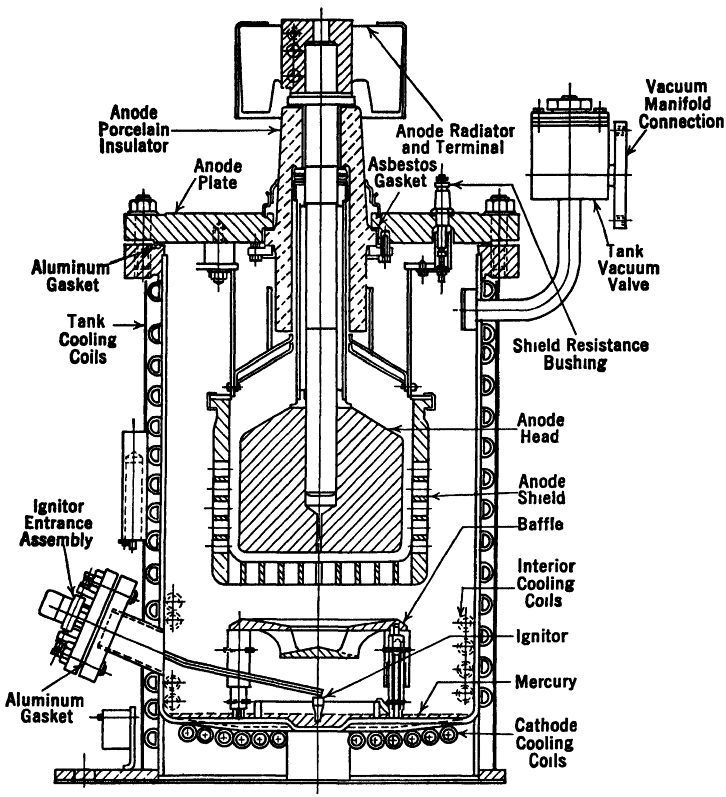


FIG. 33. Cross section of a large-capacity ignitron. (Courtesy Westinghouse Electric Corporation.)

timer contact. As long as the timer contact is "open," neither ignitor can be fired and the primary circuit is "open," being blocked by the cathode-anode gaps in the two ignitrons. The welding operation is started by the closing of the timer contact which permits current to pass through the two ignitors in series along the path indicated by the arrows. This current produces a spark at both ignitors and causes that tube having a positive anode to fire. After firing, the potential across this ignitron and the two ignitors in turn falls to a low value,

causing the current through the ignitors to drop to zero. As the a-c voltage rises on the next half-cycle, a flash of current will again pass through both ignitors and the other ignitron will fire. Thus the two ignitrons serve as an "electronic switch" to pass the alternating current through the transformer primary for operating the welder. Whenever the timer contact is opened the ignitrons again block the circuit.

Ignitron versus Thyratron. Although ignitrons and thyratrons have many characteristics in common, they are not competitive in their applications. The thyratron is essentially a low-current and high-voltage

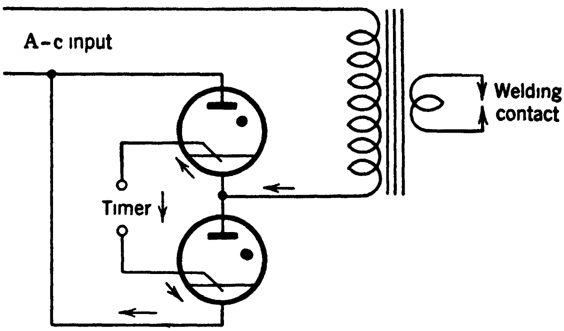


FIG. 34. Circuit for an electronic control switch consisting of two ignitrons.

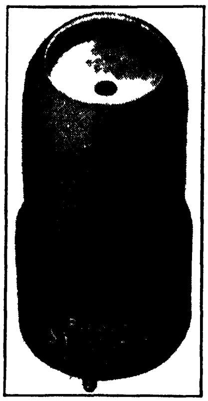
device. It can be made in small sizes for many small power applications. It has the advantage of greater simplicity of control circuits and the disadvantages of (1) the continuous use of energy for heating the cathode when it is not in use, and (2) the likelihood of damage from overload and overvoltage.

The ignitron is essentially a large-current and a low- or medium-voltage device. It is suitable for those applications having high peak current and high power requirements where the cost of the complicated ignition system does not form an excessive percentage of the total cost. The ignitron has the advantage of exceptional sturdiness, and it will withstand temporary heavy overloads and even short circuits which would ruin a thyratron. The ignitron does not require any energy for operating the cathode except that used for the ignitor when the tube is supplying rectified current. The energy for operating the ignitor is very small since the actual firing current flows for a few microseconds.

Capacitron. Another method of "firing" a mercury-pool rectifier uses an insulated conductor placed above the mercury pool. Either a

high-frequency field or a high-voltage surge impressed on this conductor will serve to control the time and firing of a pool tube. A device using this principle is called the *capacitron*. The device has little application at present.

Cold-Cathode Tubes. A number of tubes using a cold metal (other than mercury) for cathodes have been developed and employed for useful purposes. These tubes are filled with inert gas under low pressure and operate on the principle of the glow discharge. They serve (1)



Classification—recording lamp (cold cathode)	
Overall length, max	3¼ in.
Max diameter	1.275 in.
Mounting position—any	
Starting voltage, max	170 d-c volts
Operating voltage, max	135 d-c volts
Operating current	5–35 ma
Frequency range	15–15,000 cps
Useful light range	3500–6500 A

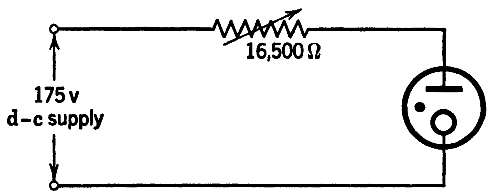


FIG. 35. Cold-cathode variable-light tube and circuit (Courtesy Sylvania Electric Products, Inc.)

as sources of light, (2) for voltage regulation, (3) for rectification, and (4) as "trigger" or control tubes. The *glow tube* has two cold electrodes and gives forth light (negative glow) when the breakdown voltage is applied (see Fig. 8, page 108). If a d-c voltage is applied the glow appears on the surface of the cathode, but if a-c voltage is impressed the glow appears at both electrodes since each serves alternately as a cathode. With simple flat electrodes lying in the same plane and with a constant source of potential the tube serves as a constant source of light and may be used as a signal light, pilot light, and test light. The source of light between two electrodes may be varied by changing the geometry of the electrode construction. Thus if the cathode is a small solid cylinder closely surrounded by a hollow cylinder, the light appears at the ends of the electrodes. The intensity of the light produced will vary as the value of the current conduction. Thus with proper construction a glow tube may become a modulated light source. Such tubes have been used for recording sound on film (light on sound track), for early types of television receivers, and for

facsimile transmission today. A commercial tube of this type, its rating, and its circuit are illustrated in Fig. 35.

For high-speed photography and stroboscopic work special types of glow tubes have been developed. One of these tubes (Fig. 36) utilizes

Classification—Strobotron cold cathode	
Physical	
Overall length	4 9/32 in.
Seated height, max	3 5/8 in.
Diameter, max	1 3/16 in.
Mounting position—any	
Electrical design characteristics	
Anode voltage	350 volts, d.c.
Average anode current, max	50 ma
Instantaneous anode current, min	5 amps
Grid No 1 d-c voltage, max	70 volts
Grid No 2 d-c voltage, max	70 volts
Grid current (max average either grid)	15 ma
Electrical Operating Characteristics	
Grid No 2—cathode starting voltage	80–125 volts
Pulse frequency range	240 pps
Tube voltage drop	
Glow discharge	75 volts
Arc discharge	20 volts

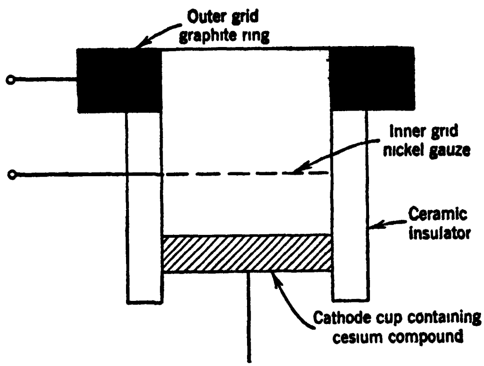
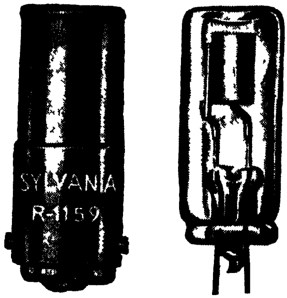


FIG 36 Stroboscopic cold-cathode light source. (Courtesy Sylvania Electric Products, Inc)

a double-grid structure which permits quick firing and gives brilliant flashes lasting for only a few microseconds. This tube was developed by Germeshausen and Edgerton. The cathode cup contains a cesium compound that liberates cesium at a relatively low temperature. Breakdown is initiated between two of the electrodes, grid to cathode, or grid to grid. Application of this tube is covered in Chapter XXII.

One important characteristic of the glow-discharge tube is the nearly constant voltage drop for a wide range in current variation. This property is utilized on a form of glow tube known as a *voltage-regula-*

tor tube. A commercial regulator tube of this type is illustrated in Fig. 37. Larger tubes having constant operating voltages up to 150 volts



Length	1¾ in.
Diameter	⅝ in.
Bulb—metal enclosure	
Starting voltage, max	170 d-c volts
Regulating voltage	65–75 volts
Operating current	0.5–5.0 ma
Regulation	±1.5 volts

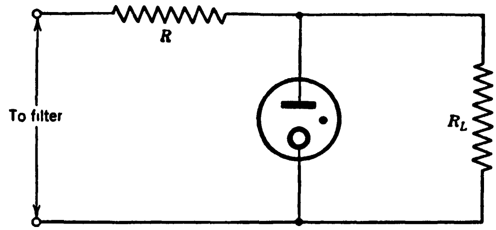


FIG. 37. Miniature voltage-regulator tube, rating, and circuit. (Courtesy Sylvania Electric Products, Inc.)

and currents up to 40 milliamperes are available. An early type of cold-cathode rectifier was known as the Raytheon BH (Fig. 38). It was a full-wave rectifier having a metal shell (an inverted metal bowl) for a cathode and two anodes consisting of insulated metal wires, the ends of which projected into the bowl. The tube was designed to utilize Paschen's law (see page 103). Thus the cathode-to-anode spacing gave the minimum voltage for ionization, whereas the short anode-to-anode spacing permitted a very high voltage for the gas pressure used. Thus conduction does not take place from anode to anode. The normal cathode-to-anode voltage drop was rather high and the operation was very sensitive to the gas pressure. Changes in gas pressure due to leaks or clean-up made the tube inoperative. Hence this device was withdrawn from the market in the early thirties and has been replaced by hot-cathode rectifiers.

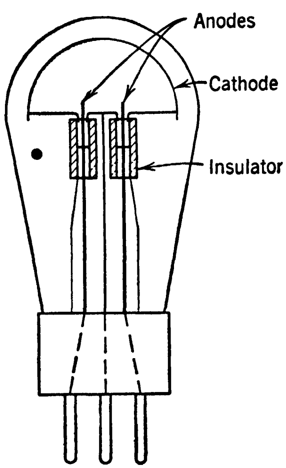


FIG. 38. Obsolete cold-cathode full-wave rectifier tube.

Another type of cold-cathode, grid-glow rectifier tube finds wide application in subscriber's telephone apparatus where its use simpli-

fies selective ringing on four-party lines. The essential parts of this tube are shown in Fig. 39. The tube employs two semicircular bell-shaped disks as cathodes. Between these cathodes is placed an anode consisting of a small circular rod enclosed in a glass tube. The tube contains neon gas under low pressure. A nominal breakdown potential of 70 volts is required to create a glow discharge from cathode to cathode, but because of greater spacing a breakdown potential of 175

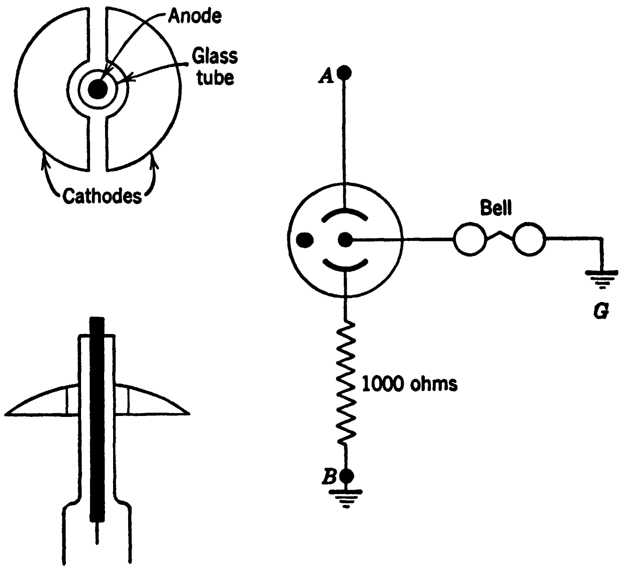


Fig. 39

volts is necessary for a like discharge from either cathode to the anode. However, while a glow discharge exists between cathodes, a potential of only 75 volts is necessary for conduction between either cathode and the anode.

This tube is used in the circuit shown on the right of Fig. 39. To ring the bell an a-c voltage is impressed across the line at points A and B, causing a glow discharge between cathodes with a minute current through the high resistance. Simultaneously, a pulsating unidirectional potential is impressed between the line A and ground G. The latter voltage causes the pulsating current to pass through the telephone bell. Individual party ringing is effected because the bells are selective with respect to the direction of current pulses and are properly distributed between the line wires and ground. One of these marginal rectifier tubes is shown in Fig. 40.

One type of gaseous rectifier starts with a cold cathode and operates on thermionic emission. The tube contains inert gas and has an oxide-coated cathode without a heater. With a minimum peak anode supply potential of 300 volts, a glow discharge takes place and the resulting bombardment of the oxide-coated cathode raises its temperature to the point necessary for satisfactory thermionic emission. After the cathode is heated the cathode-anode average potential drops to about 24 volts. This voltage is higher than that of a hot-cathode type with heater because the energy for heating the cathode is developed in the cathode-anode circuit. These tubes are known as *ionic-heated cathode* rectifiers. The output current must be limited by a suitable load impedance and a minimum cathode-anode current is required to keep the cathode hot. Examples of the ionic-heated cathodes are the OZ4 and the OZ4-G. The principle of the ionic-heated cathode is employed in the operation of some types of fluorescent lamps.

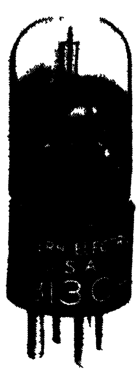


FIG. 40. Cold-cathode relay tube. (Courtesy Western Electric Company.)

The grid-glow tube * is a gas-filled, three-electrode tube having a cold cathode. The tube functions somewhat like the thyratron through the trigger action of its grid. Since it has a cold cathode the tube conduction is of the glow-discharge type involving much higher anode and grid voltages. These higher voltages mean a higher percentage of loss in the tube and lower rectification efficiency. However, the cold cathode does not consume any energy even though the tube is connected in a circuit continuously. The grid-glow tube is constructed much differently from a thyratron, as shown in Fig. 41. The anode is a wire encased in a glass tube and placed at the center of a much larger cylindrical tube which constitutes the cathode. A metal tubular shield surrounds the glass tube and anode. The grid is a small ring surrounding the projecting top of the anode. The anode is usually operated at a voltage above the breakdown voltage for the glow discharge of the gas from anode to cathode but below the breakdown for grid-to-anode spacing (see page 103). If the grid is negative or of low potential, most of the lines of the electrostatic field terminate on the grid and no breakdown from cathode to anode can occur. The usual

* The term grid-glow tube has been applied to the thyratron by some writers in the past. The usage here is in accord with current standards.

type of circuit for utilizing the grid-glow tube is shown in Fig. 42. The shield is connected to the cathode outside the tube by a resistance of the order of 5 to 10 megohms. The control grid potential is adjusted by the values of *R* and *C*. The load in the anode circuit is usually a relay. Changes in *R* or *C* or even in the value of the impressed a-c or d-c voltage will serve to fire the tube. Change in the capacity of *C* may be produced by the capacity of an approaching human hand or some other object. A photocell may be substituted for *R* or *C* whereby a change in light will fire the tube.

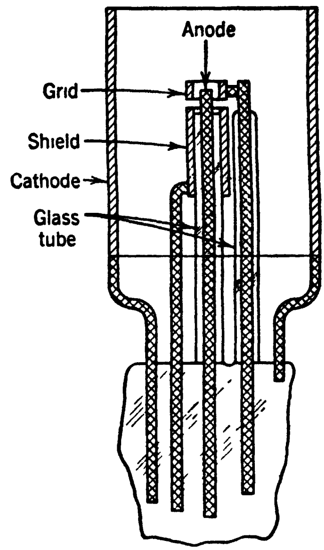


FIG. 41. Construction of a grid-glow tube.

The cathode of the grid-glow tube is indestructible but too high a voltage on the anode will cause the glow discharge to change to an arc and thus burn out the tube.

The grid-glow tube is used where a very sensitive control is desired to operate a relay and where the current delivered by the tube is sufficient to meet the requirements of an application. A commercial grid-glow tube, its rating, and its characteristics are given in Fig. 42.

Effect of Gas in a Triode. The discussion of the high-vacuum triode in Chapter IV ignored the presence of a small amount of gas which remains after the best evacuation processes are used. An interesting phenomenon produced by the presence of this gas may be observed if delicate measurements are made of the grid current under varying grid potentials and plotted to an enlarged scale. This phenomenon is illustrated in the curve of Fig. 43, which shows a reversal of the direction of grid current in the region of zero grid voltage. Some of the gas molecules in this high-vacuum triode are ionized by bombarding electrons and the positive ions thus formed are attracted to electrodes having a negative potential. When the grid is made positive, part of the electrons emitted by the cathode will be attracted to the grid and will constitute the positive grid current as shown by the curve and by the full-line arrow in the circuit diagram of Fig. 43. Now as the potential on the grid goes to zero, or negative, the grid will attract positive ions and the positive ions landing on the grid will take electrons

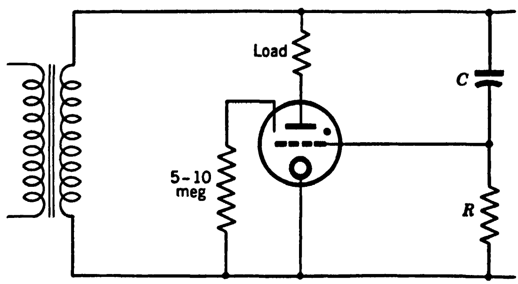
GRID-GLOW TUBE

General Characteristics

Air-cooled tetrode	
Cold cathode	
Ionization time	10 μsec
Deionization time	1000 μsec
Tube voltage drop, average	180 volts
Control characteristics—positive and negative	
Mounting position—any	

Maximum Ratings

Anode voltage, peak forward	500 volts
Anode voltage, peak inverse	800 volts
Anode current, average	0 015 amp
Anode current, peak	0 10 amp
Anode current, surge for design only	1 0 amp
Grid voltage, peak, positive	300 volts
Grid current, average	0 2 ma
Grid current, peak	0 8 ma
Averaging time, anode and grid currents	10 sec
Temperature range	−40° to +70° C



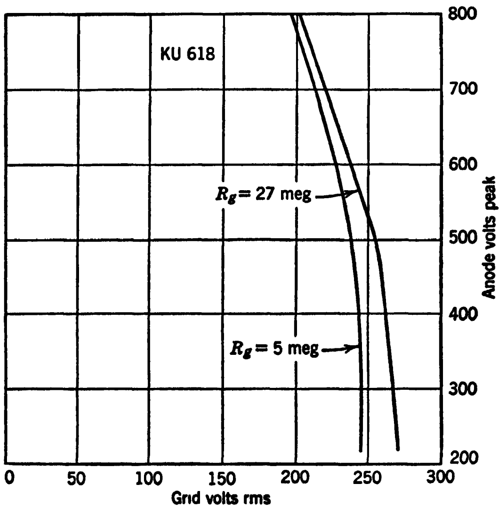


Fig. 42. Grid-glow tube characteristics, rating, and circuit (Courtesy Westinghouse Electric Corporation)

from the grid and become neutral atoms. Since the grid now has no internal source of electrons to neutralize the positive ions, electrons must come through an external circuit leading from the cathode or anode through the grid circuit to supply this need. Hence the current in the grid circuit reverses or becomes negative as shown by the curve and by the dotted arrow. As the grid is made strongly negative the electrons emitted are blocked from passing to the anode and ionization ceases. This cessation of ionization would be expected to reduce the grid current to zero, yet at this juncture the magnitude of the negative grid current begins to increase along the linear path indicated by the dotted line. This reversed current is a minute leakage current produced by the negative grid potential impressed across the high (usually tens of megohms) resistance between the glass-insulated leads to the grid and cathode. It should be noted that the reversed grid currents are very small in magnitude and are not indicated by ordinary meters. Hence this phenomenon is of theoretical interest primarily though it may assume importance where measurement of infinitesimal magnitudes are involved.

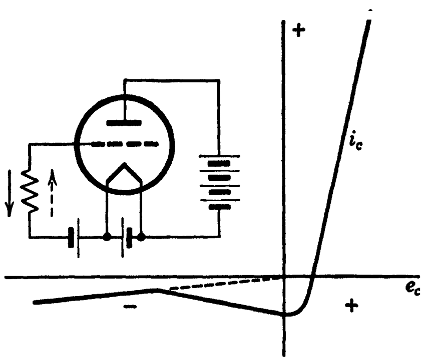


FIG. 43. Grid current in a high-vacuum triode.

PROBLEMS

1. The mercury-vapor rectifier tube of Fig. 6 is connected in series with a supply of 115 volts a-c and a load resistor. If the arc drop is assumed to be 15 volts and the peak pulse current is 0.5 ampere, what is the minimum value of load resistance to protect the tube from overload on continuous service?

2. The phanotron of Fig. 7 is connected for half-wave rectification in series with a load resistor to a 230-volt a-c supply. The rated average load current is 20 amperes which is 0.318 times the peak of the current pulse. Assume an arc drop of 15 volts and calculate (a) the magnitude of a load resistor to permit full load continuously.

(b) Now assume that, with the calculated resistor in the circuit and with the cathode only partially heated and giving a saturation emission of only 2 amperes, the supply-line switch is closed. What current will flow? What voltage exists across the cathode-anode circuit? What will happen to the phanotron?

3. (a) The thyratron of Fig. 19 is connected in series with a load resistor and 115-volt a-c supply line. What should be the value of the load resistance to limit the current to the rated average value?

(b) Assume that while the thyratron is cold and there are zero volts on the grid the supply line is connected to the cathode-anode circuit followed by a closing of the circuit for heating the filamentary cathode. What will happen?

4. The thyratron of Fig. 19 is in normal operating condition and a d-c voltage of 400 volts is to be applied to its anode. What voltage must be applied to its grid to prevent firing? Give the range of grid voltage values necessary to cover the complete range of uncertainty arising from ambient temperatures.

5. Assume that the thyratron of Fig. 19 is to be fired by an a-c voltage which rises to 30 volts + maximum and that the arc drop from cathode to grid is 10 volts. What should be the value of a resistor placed in series with the grid to limit the grid current to 0.11 ampere?

6. The thyratron of Fig. 20 is to be used for half-wave rectification of an a-c voltage peaking at 6200 volts, 60 cycles. A grid bias of -10 volts is applied to the tube. What should be the magnitude of the load resistance to limit the peak current to the rating shown?

If it is desired to fire the tube with a lag of 90 degrees behind applied voltage by superposing an in-phase a-c voltage in series with the grid (-10-volt bias), what should be the peak value of this a-c grid voltage? Use the middle curve of the characteristic. Is this particular condition for firing desirable? Discuss.

7. Construct a circuit for using the shield-grid thyratron of Fig. 21 for half-wave rectification with an anode peak voltage of 2000. Indicate the magnitude of voltages and currents at all parts of the circuit. Take values from ratings given in figure.

8. Construct a simple circuit showing how an ignitron may be fired by a thyratron using the discharge from a condenser.

9. The strobotron of Fig. 36 is flashed by the discharge of a 4-microfarad condenser (charged to 200 volts) in 100 microseconds. How much power does this represent in watts?

10. The thyratron of Fig. 19 is to be operated in a circuit having an impressed a-c sine-wave voltage of 710 rms. What negative bias voltage should be applied to the grid to cause the tube to fire with a phase-angle lag of 30 degrees, 45 degrees, 60 degrees, 90 degrees?

11. In the phase-shift circuit of Fig. 15, R has a value of 400 ohms and L of 500 millihenries. Calculate the angle of a phase shift on a 60-cycle circuit.

12. In the phase-shift circuit of Fig. 15, R has a value of 4000 ohms and C of 2 microfarads. Calculate the angle of phase shift on a 60-cycle circuit. (See text for relationship of R and C in shift circuit.)

13. What should be the extreme values of R in Problem 12 to give phase shifts of 20 degrees and 75 degrees?

REFERENCES

HULL, A. W., "Hot-Cathode Thyratrons," *Gen. Elec. Rev.*, **32** (1929).

WINOGRAD, H., "Development of Excitron-Type Rectifier," *Trans. Am. Inst. Elec. Engrs.*, **63** (1944).

SLEPIAN, J., and L. R. LUDWIG, "A New Method for Initiating the Cathode of an Arc," *Trans. Am. Inst. Elec. Engrs.*, **52**, 693–700 (1933).

GERMESHAUSEN, K. J., and H. E. EDGERTON, "A Cold-Cathode Arc-Discharge Tube," *Elec. Eng.*, **55**, 790 (1936).

Chapter VII

PHOTOELECTRICITY

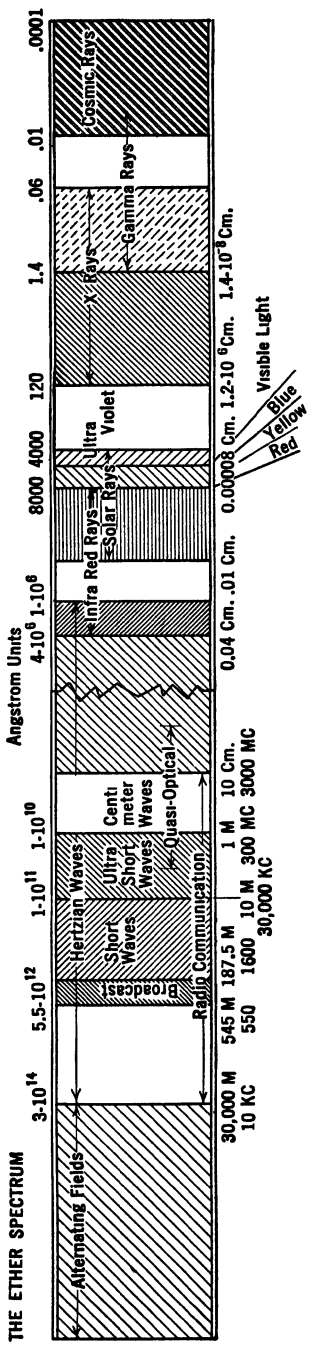
In 1905 Einstein suggested that light consists of closely packed bundles of energy which he termed "light quants." These light quants or darts of light could travel long distances and still maintain the same quantity of energy, enough to emit electrons from a solid. This theory has come to be accepted as the quantum theory of light and of radiant energy. The energy bound up in the light quant is commonly called the "photon." The concept of the scientist today regarding light is a dual theory—light is a form of wave energy and a corpuscle (photon) at the same time.

It is well to remember the reversible energy relationship in our electron theory of today. Thus in a gaseous conduction tube a flying electron may collide with a molecule of gas and give up its kinetic energy in exciting that molecule and cause it to give radiant light energy—in photons—and again a photon of light impinging upon a surface may eject an electron into space.

Spectrum of Radiant Energy. All forms of radiant energy are propagated at the velocity of light which is 3×10^{10} centimeters per second in free space. Radiant energy is assumed to travel in the form of waves and the number of waves per unit of time is called the frequency. Since the rate of propagation is fixed, it follows that

$$\text{Frequency} \times \text{wavelength} = 3 \times 10^{10}$$

This equation gives a simple relation for determining either the frequency or the wavelength when one value is known. Wavelengths may be expressed in meters, centimeters, or angstrom units. The angstrom unit is 10^{-8} centimeter or 10^{-5} micron and is commonly used for radiant energy, particularly in the visible or near-visible range. A chart showing the entire spectrum of radiant energy is given in Fig. 1. The energy in a wave depends on the frequency. Thus cosmic rays are more effective than the gamma rays which, in turn, are stronger than the X rays. The range of radiation visible to the human eye covers but a



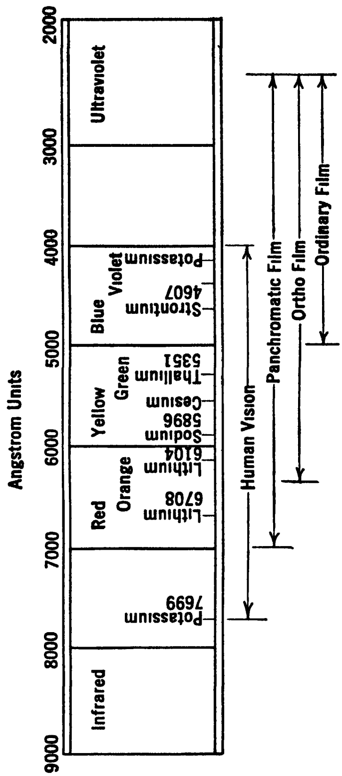


FIG 1 Spectrum of radiant energy and visible spectrum. (Courtesy *Electronics*)

very small part of the spectrum (approximately from 4000 to 8000 angstrom units). Below the visible rays are the heat rays which, in turn, lie above a wide band of radio waves. The visible spectrum is shown in the lower right-hand corner of Fig. 1.

Photoelectric Action. The term photoelectric implies the action of light in producing electricity. Since all electrical phenomena involve a displacement of electrons, the term photoelectric means the action of light upon electrons. When photons of light impinge upon solids they have the power of disturbing or releasing the electrons within that solid. The resulting movement of electrons gives rise to various kinds of electric phenomena. Thus electrons may be emitted from a solid and the action is termed *photoemission.* The electrons may be moved across a barrier in the solid, producing a difference in potential, giving rise to a *photovoltaic* action. Lastly, the photons may penetrate the solid sufficiently to release electrons and change the resistance of the substance, giving rise to the term *photoconduction.* Commercial devices have been built to utilize these three different aspects of photoelectric action.

Photoemission. In 1888 William Hallwachs found that, if he charged a zinc plate to a negative potential and then exposed it to ultraviolet light, it gradually lost its charge. However, the plate did not lose its charge when exposed to ultraviolet light after being raised to a positive potential. This phenomenon has been thoroughly investigated since the time of Hallwachs and it has been found that metals and many other substances as well exhibit it to a greater or lesser degree. This phenomenon wherein the electric charge upon a body may be changed by light is known as the Hallwachs effect.

Two famous co-workers, Julius Elster and Hans Geitel, experimented with many metals and observed that the photosensitivity of aluminum, magnesium, and zinc was superior to that of many other metals. From this they reasoned that the more electropositive metals would give better results. This expectation was proved by experiments with the alkali metals, sodium and potassium, and later with amalgams of these metals placed in an evacuated tube. Later experiments by others have shown that cesium, rubidium, and the alkaline earths, particularly strontium and barium, are satisfactory photoelectric emitters.

The theory of photoelectric emission is analogous to that of thermionic emission. Photoelectric emission results from the kinetic energy of the photon of light being imparted to the electrons in the surface of the emitter. In thermionic emission the kinetic energy of agitation due to the high temperature of the cathode gives some electrons at the

surface sufficient velocity to overcome the electron affinity at the surface. Satisfactory photoelectric emitters have a low work function. According to the Bohr theory, their atomic structure consists of a nucleus surrounded by a closely grouped band or completed system of electrons plus one lone outer electron in alkali metals, and two lone electrons in alkaline earths. It is these lone electrons which are subject to removal. The photons of light meeting these lone electrons may add to the normal velocity of the electrons sufficiently so that they overcome the electron affinity or work function and fly away from the emitting substance.

There are two laws of photoelectric emission. The first states: *The number of electrons released per unit of time at a photoelectric surface is directly proportional to the intensity of the incident light.* This law has been tested for a range of intensities varying from zero to full sunlight. Wherever apparent deviations have been discerned they can be explained by errors in measurement or by inherent faults within the cell employed which tended to introduce spurious currents or to hinder the total number of electrons actually released from being collected. This law makes the principle of photoemission very valuable for light measurement and many other applications. The second law states: *The maximum energy of electrons released at a photoelectric surface is independent of the intensity of the incident light but is directly proportional to the frequency of the light.* This law implies that the energy imparted to the electron by electromagnetic radiations is directly proportional to the frequency of these radiations.

The reason for the first law of photoemission is rather obvious but that for the second law is not so clear. The second law was explained in 1905 by Einstein who reasoned that incident radiant energy was transferred to surface electrons in a quantum of magnitude hf wherein a portion of this energy was used for removal of the electrons and the remainder appeared as kinetic energy. In mathematical form Einstein states the relationship as follows:

$$hf = \phi + \tfrac{1}{2}mv^2 \tag{1}$$

where h is Planck's constant, f is the frequency of the incident light, and ϕ is the work function of the emitter in equivalent electron-volts. The term hf is the energy of the impinging light (photons) which overcomes the work function ϕ and gives a velocity v to the emitted electrons. Inspection of equation 1 shows that if hf equals ϕ the electron will have zero velocity and will not be emitted. Thus there is a mini-

mum or *threshold frequency* for the incident radiation below which photoelectric emission will not result. Equation 1 has been verified experimentally. The second law of photoemission follows from equation 1 because h and ϕ are constants for a given emitter, making $\frac{1}{2}mv^2$ proportional to f. Equation 1 holds only for the electrons lying on the surface which will have the maximum velocity of emission. Some light will penetrate the outer layer of atoms and liberate electrons within the emitter, but these electrons are likely to lose much of if not all their energy before leaving the surface.

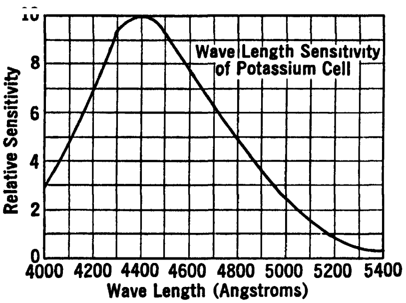


FIG. 2. Photoemissive color sensitivity of potassium. (Courtesy General Electric Company.)

Color sensitivity is the relative response of photoemissive materials to various light radiations. A curve of the sensitivity of potassium is given in Fig. 2. This curve covers the green-blue-violet part of the visible spectrum with the peak occurring in the violet range. Photosensitive materials having a peak in the red range of the spectrum are very desirable for use under the light from incandescent lamps which radiate their maximum energy in the near infrared.

Photosensitive films generally show a much greater emission than a pure metal or substance. In this respect and several others, photoelectric emission parallels thermionic emission. Thus thermionic emission from tungsten is greatly increased by a film of thorium one molecule thick, and emission from a metal covered by barium oxide with a monomolecular layer of barium on the outside is still greater. The surface treatment of thermionic emitters also greatly affects the resulting emission. These various factors control the work function of the emitting surface. In a similar manner the photoelectric emission may be improved by a thin (molecular) layer of cesium on magnesium, whereas a much greater emission may be attained by a thin deposit of cesium upon a sub-base of cesium oxide covering a base of pure silver. Again, the treatment of potassium in hydrogen will change the color sensitivity and will improve the emission of that light-sensitive material.

Vacuum Phototubes. The photoemissive tube consists of a cathode having a photosensitive surface and an anode placed in a glass envelope. These two electrodes are connected in series with a battery

(anode positive) and a load consisting of a resistance as shown in Fig. 3. Light falling upon the cathode causes the emission of electrons which are attracted to the positive anode and which cause a current through the external circuit including the load. In the vacuum phototube the space within the glass envelope is highly evacuated.

Many different types of construction have been used for phototubes. In the early type the cathode frequently consisted of a deposit of light-sensitive material on the inner wall of the spherical tube. A circular window for admitting light was made on one side of the sphere by evaporation using the heat of a Bunsen burner. Anodes usually consist of a straight wire or wire ring placed near the center of the envelope. Sodium, potassium, barium, rubidium, and cesium have been the materials used for the light-sensitive surface. Nearly all commercial tubes produced today are enclosed in cylindrical glass envelopes as shown in Fig. 4. The cathode is a plate bent into the form of a semicylinder. The light-sensitive material is deposited on the cathode after the tube is assembled and evacuated. For the measurement of ultraviolet light either sodium or cadmium may be used. In order to admit the ultraviolet light, quartz or some special glass such as Corex must be used. One scheme for admitting ultraviolet is to use a special thin circular window (1-inch diameter) of Corex glass which projects out from the standard cylindrical glass envelope. The highest sensitivity of phototubes to light from incandescent lamps is obtained through the use of the cesium oxide tube (developed by L. R. Koller) which consists of a monomolecular coating of cesium on a sub-base of cesium oxide upon a base of silver. This type of tube may be produced in a number of ways, one of which is as follows.*

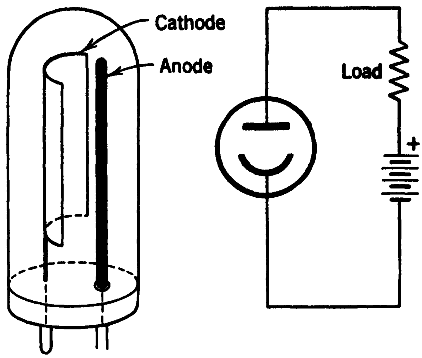
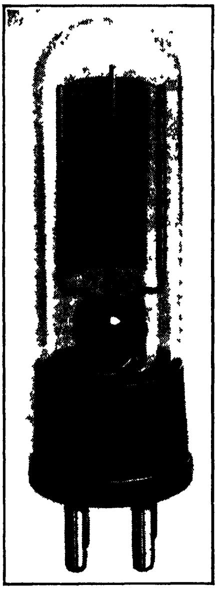


FIG. 3. Construction and circuit of a phototube.

The cathode is made of solid silver or of silver-plated copper in the form of a semicylinder. The anode is a coaxial vertical wire with a disk of nickel welded to its top. The disk holds in a pocket a small pellet of cesium chloride and calcium or of cesium dichromate and

* The following description is taken from *Photocells and Their Application,* by V. K. Zworykin and D. Wilson, John Wiley & Sons, 1934.

silicon. The glass press and its electrodes are sealed into a cylindrical glass bulb. The tube is exhausted and baked in a furnace. After cooling, about 2 millimeters of oxygen are admitted and a glow discharge is passed between the anode and cathode. Oxidation of the silver is evidenced by a progressive series of brilliant interference colors on the cathode. The glow is continued intermittently until the color of the cathode passes through the first bright green. The residual



General Characteristics		
Vacuum and gas diodes	Vacuum	Gas
Cathode surface	S_1	S_1
Cathode size	0.81 x 1.38	0.81 x 1.38 in.
Luminous sensitivity, 0 cycles, μamp/lumen	15	50
Luminous sensitivity, 10,000 cycles, μamp/lumen	15	55
Radiant response μa/μW at 7500 Au	0.0015	0.0060
Response max at	7,500	7,500 A
Response, upper limit	11,000	11,000 A
Response, lower limit	4,000	4,000 A
Gas amplification, max	.	7
Capacitance, anode-cathode	2 5	2 5 μμf
Leakage resistance, min	4,000	90 meg
Mounting position	Any	Any
Net weight	2 oz	2 oz
Shipping weight	8 oz	8 oz
Maximum Ratings		
Anode voltage, peak	500	90 volts
Anode current	20	20 μamp
Temperature, ambient air	100°	100° C

FIG. 4. Phototube and characteristics. (Courtesy Westinghouse Electric Corporation.)

gas is then pumped out. By means of a high-frequency field, the pellet disk is heated to a temperature high enough to explode the pellet and liberate cesium vapor which initially condenses on the glass walls of the bulb. The tube is then placed in a furnace and held at a temperature of 200 to 225 degrees C. If the back of the cathode or the stem of the press is painted with a mixture of lead oxide or tin oxide in amyl acetate, all excess cesium will be absorbed by the paint and the cathode will become dark gray in color as the baking proceeds. When the edges of the cathode begin to turn light in color, the tube is completed. The color-sensitivity curve of a cesium oxide vacuum phototube made in this manner is given in curve S_1 of Fig. 5. The sensitivity of this tube is of the order of 10 to 20 microamperes per lumen.

Since the current output of vacuum tubes is so small, any leakage between the cathode and anode lead-in wires may be troublesome. For low light intensities any leakage current through the glass or on

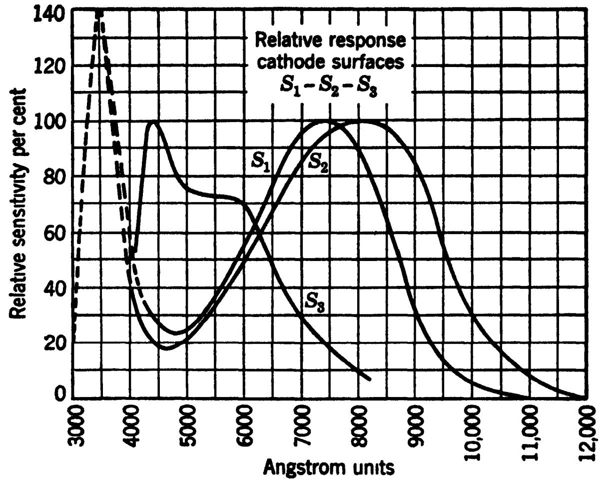


FIG 5 Relative responses of three different photosensitive surfaces (dashed lines are outside the range of human vision) S_1 is caesium-oxygen-silver, S_3 is rubidium-oxygen-silver, and S_2 is discontinued (Courtesy Westinghouse Electric Corporation)

the inside or outside walls of the envelope may mask the real emission current The magnitude of the leakage current may be reduced in two ways If the leads come to the tube base, a conducting ring may be sealed in the glass stem around the anode lead and connected in shunt around the load resistance as shown in Fig 6. A simpler arrangement is to bring the lead of one electrode out at the base and the other out of the top of the tube, making a long path for any possible leakage current.

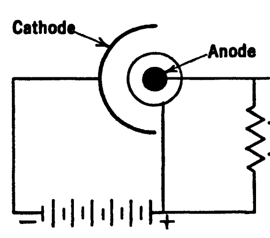


FIG 6. Circuit for eliminating leakage current in the load of a phototube.

Negative space charge has little influence in the vacuum phototube because the cathode is made relatively large to collect light and the emission is small Some space charge may exist near the anode but if this electrode is maintained strongly positive, the space charge has no detrimental effect.

The variation of the anode current with anode potential for constant light intensity and a given color value is shown by Fig. 7. This indicates that a low potential of 20 to 40 volts will attract nearly all elec-

trons emitted. Potentials of 90 to 250 volts are used in the circuits of vacuum phototubes. The resulting cathode-anode voltage drops for four values of load resistance may be obtained at the intersections of the load lines with the characteristic curve.* The curve for anode current versus light flux is a straight line following the first law of photoemission. Any deviation from this characteristic is due to fatigue in the light-sensitive material or some variable electric leakage.

The vacuum phototube is suited for applications where the use of a

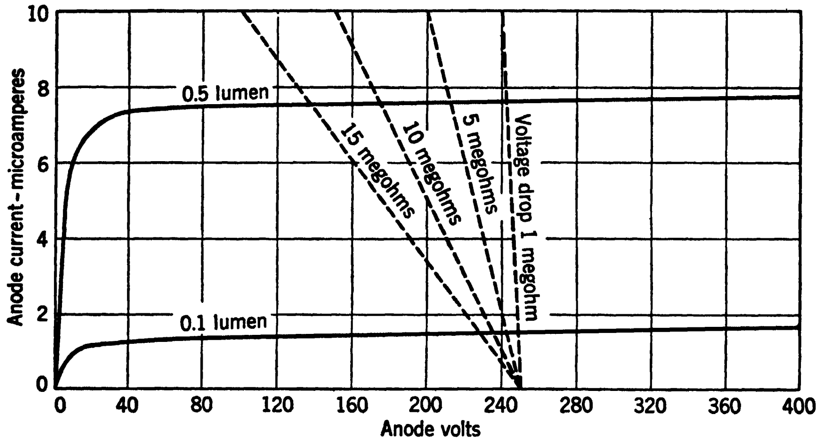


FIG. 7. Anode-currency voltage characteristic of a vacuum phototube (Courtesy Westinghouse Electric Corporation.)

high resistance load is desirable to give maximum circuit sensitivity. It offers a high stability of operation and permanence of calibration. Thus it is used for light-measuring and light-relay devices. This tube is also used where good dynamic response † is needed.

Gaseous Phototube. The admission of a small amount of inert gas into a vacuum phototube will greatly increase the current flow owing to light falling upon the cathode. This amplification of current or increase in sensitivity may be tenfold in value, but it is usually limited to the order of three to seven times the current in a vacuum. This increase of current parallels that found in thermionic tubes containing gas, but the cause of the phenomenon is somewhat different. In the thermionic tube the increase is due to the neutralization of negative

* The load line represents the voltage drop across the load resistor of Fig. 3. Subtracting this drop from the d-c supply voltage gives the voltage remaining across the phototube.

† See page 168 for explanation of dynamic response.

space charge near the cathode, whereas space charge is of little consequence in the phototube. The reasons for the current increase in the gaseous tube are (1) the additional ionic current which results as soon as the ionizing potential is reached; and (2) an increase in emission of electrons from the cathode due to the bombardment by positive ions. The change in anode current with rise in anode potential is given in Fig. 8 for five different amounts of light flux in lumens falling on the cathode. Obviously, a low anode potential attracts all the primary

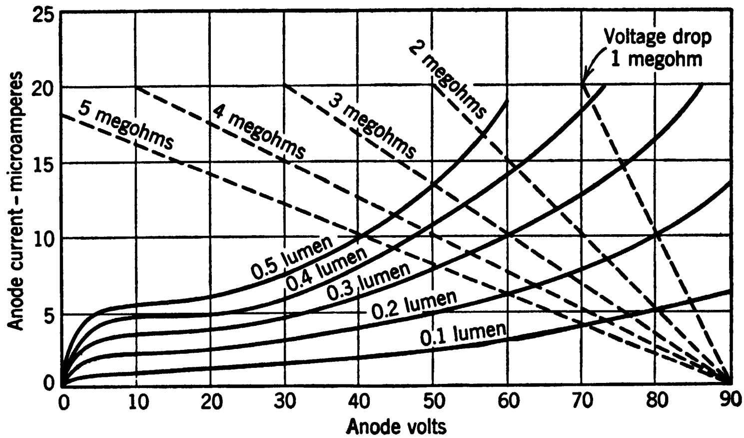


FIG. 8. Anode-current voltage characteristic of a gaseous phototube. (Courtesy Westinghouse Electric Corporation.)

electrons emitted. At 20 volts (approximately) the ionizing potential is reached and beyond that the current increases with voltage, at first slowly, later becoming accumulative in effect following the theory of the Townsend discharge described in the chapter on gaseous conduction. The potential applied to the gaseous phototube circuit is 90 volts and the actual cathode-anode drop is limited by the load resistance. The trend of the curve will depend on the gas pressure and other factors. The operation of the gaseous tube must be kept below the glow-discharge point or the control of its operation by incident light is lost.

The light-flux-current characteristic of the gaseous tube is given in Fig. 9. This curve is not linear but the deviation does not cause any undesirable distortion of music or voice when the tube is used for sound on film.

One disadvantage of the gaseous phototube is its poor dynamic response. Dynamic response corresponds to the high-frequency response

of amplifier tubes and circuits. It refers to the speed with which the anode current follows changes in the light incident on the cathode. It is important when there is a high frequency in the light changes. The poor dynamic response is due to the current conduction by ions in this cell. At a given instant the space between the cathode and anode contains a large number of electrons and positive ions. The electrons are swept out of the space instantly by the positive potential on the

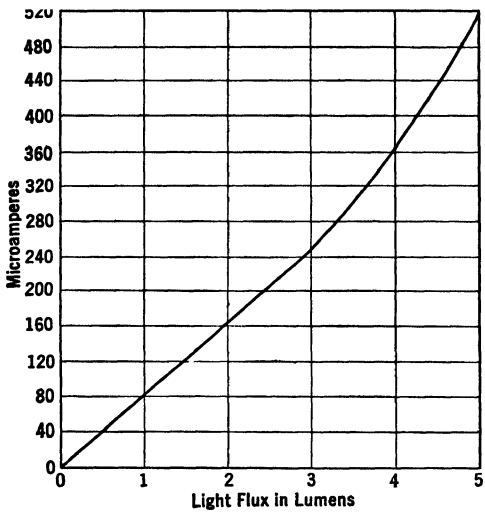


FIG. 9. Light-flux-current characteristic of a gaseous photocell.

anode but the positive ions, being relatively ponderous, move much more slowly toward the cathode. If the light drops to zero at a given instant, the positive ions take a little interval of time to reach the cathode; when they do reach it they may release a few electrons which then dart to the anode. Thus, although the light goes to zero instantly, the anode current lags in making the change. If alternating changes in light flux take place, the anode current will fall behind in making like changes and the lag will be proportional to the frequency of that change. The difference in the dynamic response between the vacuum unit and the gaseous unit is illustrated in the curve of Fig. 10.

The standard commercial gaseous phototubes use a cesium oxide cathode. They are filled with argon at a pressure of about 100 microns. (A micron is a pressure equivalent to a column of mercury one-millionth of a meter high.) They have a sensitivity of about 50 microamperes per lumen though some are built with sensitivity of 60 to 100.

The more common commercial tubes are about $1\frac{1}{8}$ inches in diameter and $4\frac{1}{8}$ inches high (including base and pins). The cathodes are about $1\frac{3}{8}$ inches high with a diameter of $\frac{11}{16}$ inch.

Curves of the color sensitivity of three types of cathode surfaces are given in Fig. 5. Surface S_1 covers the spectrum of the incandescent lamp satisfactorily. Surface S_2 is adapted for infrared rays and surface S_3 works best in the blue and ultraviolet region of the spectrum.

Gaseous phototubes are used almost universally for sound reproduction from films. They serve well in many other applications not requiring a high dynamic response. Some of these applications are: (1) for facsimile transmission of pictures, (2) as an "electric eye" in photometric measurements, (3) as an "electric eye" in color analysis and color matching, (4) for calibration of watthour meters, (5) for miscellaneous counting operations, (6) for detection of defects in manufactured articles, (7) for control of sign lighting, and (8) for control of artificial illumination in factories.

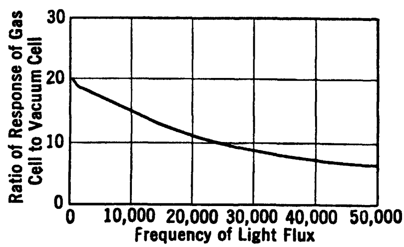


FIG. 10. Curve showing effect of dynamic response. (Courtesy Zworykin and Wilson.)

Photovoltaic Devices. A photovoltaic device is one in which the energy in incident light creates a difference in potential. In 1839 E. Becquerel discovered that, when one of two electrodes immersed in an electrolyte is illuminated, a difference in potential appears between the electrodes. This difference disappears when the electrode is in darkness. This principle has been employed in a commercial photovoltaic cell known as Rayfoto. The cell consisted of a sensitive electrode of cuprous oxide and an anode of lead immersed in an electrolyte of lead nitrate. The sensitivity of this cell was about 150 microamperes per lumen. This cell and other electrolytic types of photovoltaic cells have been displaced by the dry type of device.

In 1876, while experimenting with the conductivity of amorphous selenium rods embedded in iron, Adams and Day discovered that a difference of potential was created when light fell on their apparatus. This phenomenon of the creation of a potential difference by light falling on a junction of selenium and iron was rediscovered by Charles Fritts in 1884. Forty-six years later (1930), B. Lange observed the phenomenon for the third time.

The photovoltaic devices of the dry type consist of elements similar to those used in the blocking-layer type of rectifier—copper oxide on copper and selenium on iron. Curiously, when the device is used as a photovoltaic device the direction of the electron flow in these elements may be opposite to that observed when it is used as a rectifier. The action of a simple photovoltaic cell is illustrated in Fig. 11 (part *a*). Light falling upon a thin semiconductor layer (copper oxide or selenium) causes electrons to move to the metal (copper or iron). The elec-

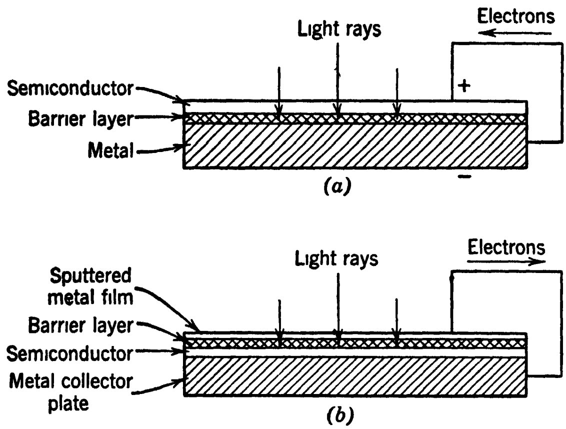


FIG. 11. Construction of photovoltaic cells: (*a*) back plate type; (*b*) front plate type.

trons that move to the metal are trapped, and they build up a negative potential in the metal. In the meantime, the semiconductor has lost electrons and assumes a positive potential, and a difference of potential is built up between the two electrodes. This difference of potential does not become very high (about 0.3 volt) because the electrons can leak back to the semiconductor through the barrier layer.

The theory of the photovoltaic cell lies in the action taking place in the barrier between the metal and the semiconductor. The semiconductor is in crystal form, and the crystals may touch the metal at points so that the area of contact is small. These crystals have a rectifying action and offer a high resistance to current flow across the barrier in one direction, but a low resistance for the opposite direction of flow. Light falling on the semiconductor penetrates it as far as the barrier crystals and there gives up the energy of the photons, which is sufficient to overcome the high resistance of the barrier and to cause electrons to move into the metal and become trapped. One may

wonder why the displaced electrons remain entrapped since the normal direction of electron movement via rectifying action is *from* metal to semiconductor. The answer to this question is to be found in Fig. 8 in the chapter on crystal and metallic rectifiers. From this figure it will be noted that, while the resistance to current in the forward direction is approximately zero for positive potentials of 1 or more volts, the resistance at low potentials of zero to 0.3 volt lies in the range of 400 to 150 ohms. Higher values may exist for cells of different dimensions and construction. Hence it is obvious that, if a low-resistance circuit or load is connected to the two electrodes of the photovoltaic

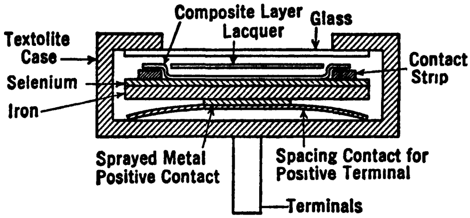


FIG. 12. Construction of a photovoltaic cell. (Courtesy General Electric Company.)

element, the leakage current can be kept small. On the contrary, a high-resistance load will result in considerable leakage.

Barrier-Layer Photovoltaic Cells. The photovoltaic devices in use today differ from the one just described in having a thin metallic layer sputtered on the surface of the semiconductor (Fig. 11*b*). Gold and platinum have been used for this surface layer. The theory of action of this cell is the same as previously explained except that the direction of electron drift is reversed. This reversal results because the active metal or metallic layer is now exposed to the light which penetrates to the semiconductor layer and gives up its energy, thus moving the electrons from the semiconductor to the surface layer. This action reverses the polarity of the cell.

It is probable that some light entering the cell may penetrate the semiconductor layer and cause a second action at the lower barrier, thus moving some electrons into the bottom collector plate. This second action will produce a counter emf effect but it will be small in magnitude.

A commercial photovoltaic cell is shown in Fig. 12. Fundamentally, the device consists of a block of iron covered by a layer of selenium. A transparent composite layer is sputtered upon the selenium surface,

after which the entire unit is covered with a film of lacquer to protect it from the atmosphere. Contacts are established with the composite layer for one electrode and with the iron back for the other electrode. When light falls on the selenium, electrons pass to the composite surface layer.

The current output of the selenium-iron device depends upon the illumination and the resistance of the external load. The variation of

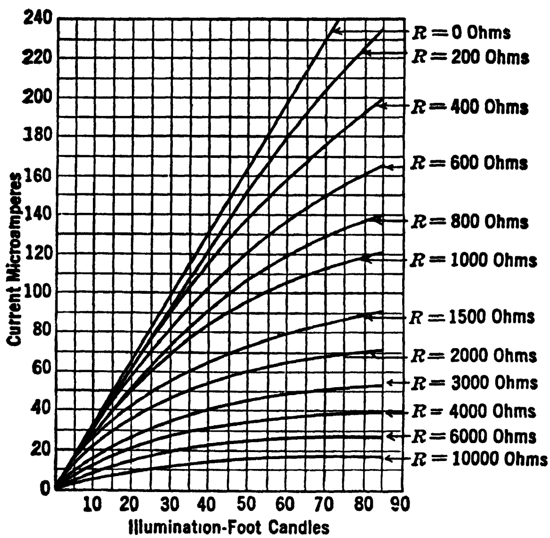


FIG. 13. Current-illumination characteristic of a photovoltaic cell for various load resistances. (Courtesy General Electric Company.)

current with these two factors is shown in Fig. 13. This variation is almost linear for zero external resistance and indicates that a low load resistance is very desirable for light-measuring devices. The change from linearity for high load resistance is due to the increase in internal leakage across the barrier. The color sensitivity of the selenium-iron cell illustrated in Fig. 12 is given in Fig. 14. The sensitivity curve compares fairly satisfactorily with that of the human eye in the range of visible light.

The characteristic of the selenium-iron cell changes with temperature, particularly where the load resistance is high. These cells should not be used at temperatures higher than 50 degrees C. The action of this device is subject to fatigue; however, the error with exposure is slight if the external load resistance is kept small. This device is also

subject to aging, but the commercial products can be aged artificially so that there is no further change in the hands of the purchaser.

The principal application of the selenium-iron photovoltaic cell has been for the measurement of light. One application known as a foot-candle meter (Fig. 15) consists of a selenium light cell and a sensitive ammeter built into a compact unit. A similar device, called an exposure meter, for photography, is illustrated in Fig. 16.

Another commercial form of selenium-barrier-type photovoltaic cell is illustrated in Fig. 17. This device is called a *photronic* cell and is supplied in three grades of sensitivity varying from 2.5 to 4.5 microamperes per foot-candle. When equipped with a viscor filter the sensitivity of these cells very closely matches that of the human eye. Photronic cells are used in circuits with sensitive d-c meters for light measurement, or they may be used to operate sensitive relays.

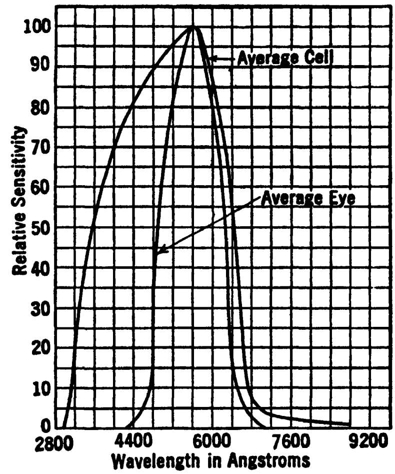


FIG. 14. Color sensitivity of a selenium photovoltaic cell.

The selenium photovoltaic cell may be connected to a meter in a simple series circuit if the resistance of the cell is suitable for damping the meter and for giving a linear deflection to the light response. If these factors are not satisfactory, shunt resistances may be employed across the meter and cell. The response of the photovoltaic cells to light change is satisfactory for light measurements, but for rapidly changing light, such as sound on film, the dynamic response is unsatisfactory.

Copper oxide photovoltaic cells have a history and a theory of operation that parallel those of the selenium-iron type. The photovoltaic action of copper oxide on copper was observed by Pfund in 1916. An English patent was obtained on the copper oxide element in 1928, and about 1935 a commercial photovoltaic cell using the copper oxide unit was placed on the American market.

A comparison of the copper oxide and the selenium-iron photovoltaic cell shows that both units are subject to effects of temperature and to aging. The important difference is in the degree of aging and the period

of aging. Here, as in the dry-type rectifier units, the copper oxide unit is subject to a greater effect of aging. Although this aging effect is not so important in a rectifier, it becomes very serious if the unit is used

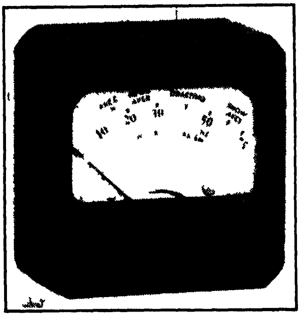


FIG. 15 Photovoltaic foot-candle meter (Courtesy General Electric Company)

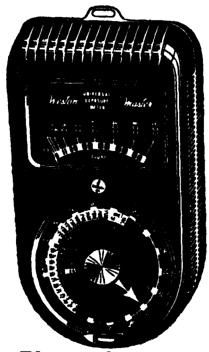


FIG. 16. Photovoltaic exposure meter. (Courtesy Weston Electrical Instrument Corporation)

for the measurement of light. The decrease in the electron emission due to aging and to high temperatures causes the copper oxide light meter to lose its calibration, and accordingly the copper oxide photovoltaic unit has been withdrawn from the market.

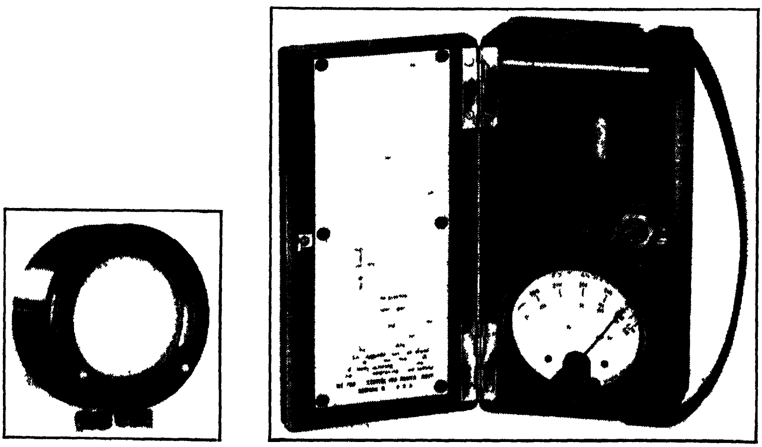


FIG. 17. Photovoltaic cell (*left*); built into a foot-candle meter (*right*). (Courtesy Weston Electrical Instrument Corporation)

Photoconductive Devices. In 1873 Willoughby Smith observed that high-resistance elements consisting of tiny rods of selenium became better conductors when exposed to daylight or to any artificial illumi-

nation. Selenium is a chemical element which lies on the borderland between conductors and insulators. Under any condition it is a poor conductor, but in the gray crystalline form it shows a decrease in resistance under the influence of light. A large number of different forms of selenium cells have been devised to take advantage of this property of selenium. Many of these cells are old in the art so that this device antedates all other photoelectric cells in time of practical use.

In order to utilize the peculiar property of gray crystalline selenium, it must be connected as in the circuit shown in Fig. 18. When light falls on the selenium its resistance decreases and the current in the circuit rises. If R represents a marginal relay it will be operated by the increase of current due to incident light. The decrease of resistance is due to the freeing of electrons by the incident light. Apparently the energy in the photons of the incident light is imparted to some of the electrons in the outer orbits of selenium atoms so as to free them from the bonds of the nuclei. Then the potential gradient in the selenium causes these electrons to move on with greater freedom. As the electrons move on they recombine with positive ions and become neutral atoms. As long as light falls on the selenium new electrons are released, but when light ceases and recombinations are completed the electron movement falls to a low value.

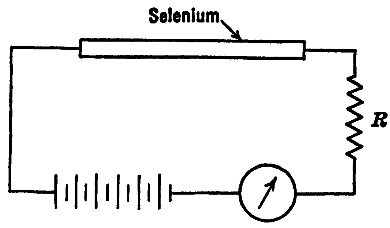


FIG. 18. Circuit for a photoconductive cell.

Since light does not penetrate the selenium very deeply and since the resistance of selenium is very high, two things are necessary in the production of a satisfactory selenium cell. First, the selenium should be used in a very thin layer to permit light penetration, and, second, the cross-sectional area for the current flow should be made as large as convenient. To accomplish these two aims many selenium cells have been made as illustrated in Fig. 19. A metal film such as gold leaf is placed on a plate of glass or other insulating material. The film is scratched by a zigzag line of grid type. Then a thin layer of selenium is placed over the whole plate and given a heat treatment to change the selenium to the gray crystalline form. After this treatment the plate may be sealed in a glass tube to protect it from the atmosphere. It will be observed that the element consists of two electrodes separated by a long, narrow bridge of selenium. When light falls on the active

side of the plate the resistance of the entire volume of selenium in the bridge is changed and a relatively large change in current may be obtained. Commercial selenium cells have shown a current ratio from light to dark as high as 4 under an illumination of 10 foot-candles.

In earlier times selenium cells were used for control operation such as turning on or off artificial light with the coming or passing of daylight. They were improved sufficiently in recent years to make them satisfactory for sound reproduction from film but did not receive much recognition. If its characteristics had been improved at an earlier date, the selenium photoconductive cell might have had a wide application.

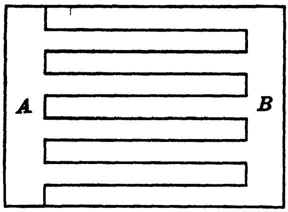


Fig. 19. Construction of a photoconductive cell.

Comparison of Photoelectric Devices. In passing, it is well to note that the photoemissive and the photoconductive devices require an outside source of d-c supply for their operation, whereas the photovoltaic device gets its energy direct from the light. The photoemissive device produces a very small current and an amplifier is needed before any useful work can be performed. The

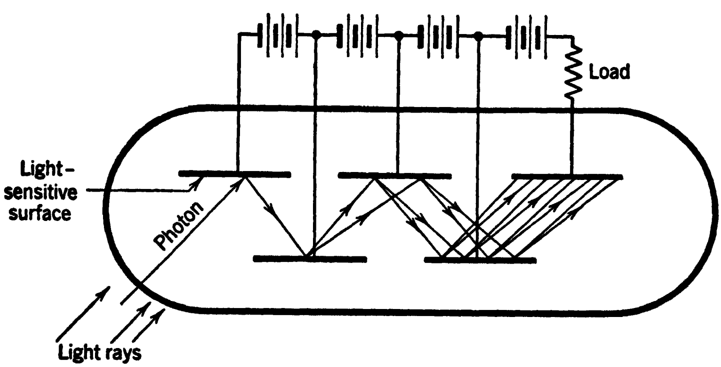


Fig. 20. Principle of a photoelectric electron-multiplier tube.

photovoltaic and photoconductive devices may operate a sensitive relay directly without any amplification. The photoemissive devices furnish a signal accurately proportional to the incident illumination and the voltage of their output is satisfactory for amplification. The photovoltaic device feeding into a low load resistance produces a current proportional to the incident illumination. When this device feeds

into a high-resistance load the voltage produced is not linear with the incident light. This fact plus the low magnitude of its voltage makes the device unsuitable for the grid supply of an amplifier.

Electron-Multiplier Tubes. Photoelectric emission is utilized for amplification of a light signal in a tube known as an electron multiplier. The theory, circuit, and one schematic type of construction of the de-

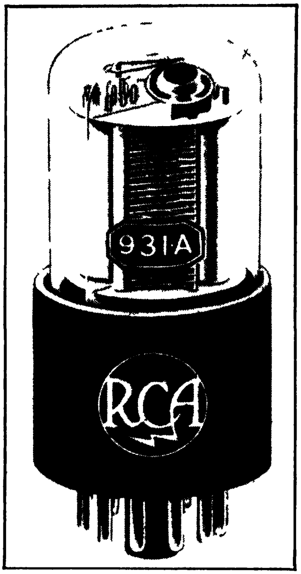


Fig. 21. Electron-multiplier tube. (Courtesy Radio Corporation of America.)

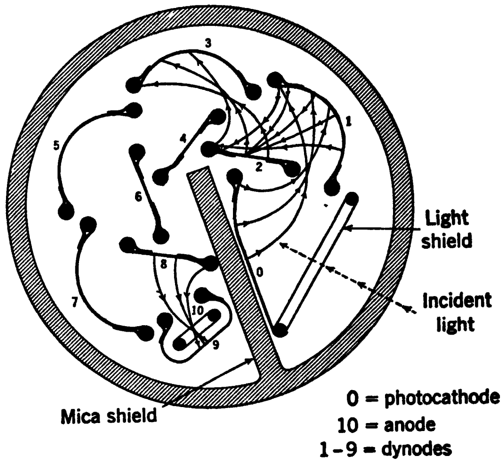


Fig. 22. Cross section of the electron-multiplier tube in Fig. 21.

vice are illustrated in Fig. 20. The first of a series of electrodes in a vacuum chamber has a light sensitive film which emits electrons in proportion to the incident light. The emitted electrons are attracted to a second electrode held at a higher positive potential sufficient to cause secondary emission. These electrons of secondary emission are drawn to a third electrode where an increased number of secondary electrons are emitted. One primary bombarding electron may emit from 1 to 10 secondary electrons, depending on the electric field and work function of the electrode surface. Thus through a series of electrodes, each having a higher potential than the preceding one, any amount of amplification may be attained theoretically. A practical limit in multiplication arises through space-charge effects and power dissipation in the final stages.

A commercial form of electron-multiplier tube is shown in Fig. 21 and an enlarged cross section of the device in Fig. 22. In this device incident light entering the tube releases electrons from the light-sensitive cathode O from whence they are drawn to electrode 1 where secondary emission begins and continues at increasing values throughout nine successive stages. Development engineers worked out an ingenious design for incorporating several anodes in a very compact space, the tube being only 1⅛ inches in diameter. With the use of relatively high potentials it is possible to attain an amplification of one million, though such a value is not attained in commercial usage.

The electron multiplier can be used as a combined photoelectric pick-up and amplifier for minute light variations, for sound on film, and in television.

Problems

1. The vacuum phototube of Fig. 4 is connected in the circuit of Fig. 3 with an applied d-c potential of 250 volts and a resistance of 10 megohms. Let the light falling on the cathode of the tube change from zero to 0.1 lumen and 0.5 lumen. Calculate the voltage drop across the resistor for each case. (Use curves of Fig. 7 for this tube.) Does the answer check with law No. 1 for photoelectric emission?

2. Connect the gaseous phototube of Fig. 4 into the circuit of Fig. 3 using a 90-volt battery and a 5-megohm resistor. Assume a light flux of 0, 0.1, 0.2, 0.3, 0.4, and 0.5 lumen falling on the cathode of the tube, and determine the voltage drop across the resistor for each case using the curves of Fig. 8. Plot curve of voltage drop versus lumens.

3. Repeat Problem 2 using a 1-megohm resistor. Compare the results of Problems 2 and 3. Which of the above tubes should be used for accurate light measurement?

4. Design a circuit for firing a thyratron using the increasing light falling on a phototube.

5. Refer to Figs. 1 and 5 and select a phototube surface for use with (a) ultraviolet light, (b) infrared light, and (c) incandescent lamp light.

6. Develop the following equations for the threshold frequency and threshold wavelength in terms of the work function in equivalent electron volts.

$$f = 2.415 \times 10^{14}\,\phi \text{ (work function)}$$

$$\text{angstroms} = \frac{12{,}400}{\phi \text{ (work function)}}$$

Hint: Use equation 1 of this chapter, equation 10 of Chapter I, and the value for Planck's constant

$$h = 6.624 \times 10^{-34} \text{ joule per second}$$

7. Calculate the threshold frequency for copper, nickel, zinc, and barium.
8. Calculate the threshold wavelength for sodium, potassium, and cesium.

REFERENCES

ZWORYKIN, V. K., and E. D. WILSON, *Photocells and Their Application,* John Wiley & Sons, 1934.

LANGE, BRUNO, *Photo-Elements and Their Application,* Reinhold Publishing Corp., 1938.

Chapter VIII

CRYSTAL AND METALLIC RECTIFIERS

The phenomenon of nonlinear voltage-current relations in some substances and at the junctions of dissimilar substances has been known for a long time. As early as 1834, Faraday discovered the highly negative temperature coefficient of resistivity exhibited by silver sulphide. Early in this century many materials were employed in "wireless" reception which exhibited highly polarized nonlinear characteristics. These materials consisted of a lump of some mineral such as galena, silicon, iron or copper pyrites, zincite, bornite, or silicon carbide. In the process of signal detection these materials made up one side of a circuit while a feeler wire called a "catwhisker" bearing upon the material was connected to the other terminal. This combination known as a crystal detector had a unilateral current characteristic in which electrons moved from the crystal to the metal point bearing upon them. The disadvantages of this early device were that the crystal had "sensitive spots" which were destroyed by overload and heat and the sensitivity could be lost by slight displacements of the catwhisker. This form of crystal detector or rectifier was widely used in the amateur period of wireless telegraphy and the early days of voice radio. The more stable operation of the vacuum-diode detector caused the crystal detector to be displaced and fall into disuse shortly after 1920. Experimentation on the construction and use of crystals as rectifiers was resumed during the 1930's, and two new crystal rectifiers, silicon and germanium, developed during this period were put into effective use during World War II.

The unilateral current-carrying characteristic of a junction of selenium and a metal was discovered in 1883, though commercial application of the combination was not made until 1930. A similar discovery of the unilateral characteristic of copper oxide on copper was made about 1920 and resulted in an early application of this combination.

The crystal and the semiconductor rectifiers suggested in the preceding statements are not electronic devices in the narrow sense of that term since they do not involve movements of electrons in a vacuum or

ions in gases under low pressure, but they do serve as important components in electronic combinations and circuits.

Silicon and Germanium Crystal Rectifiers. The basic construction of the modern crystal rectifier is illustrated in Fig. 1. A tungsten wire having an S or a hook shape has a sharp point which bears on a semiconductor or rectifying material. The tungsten wire is fastened to one terminal and the semiconductor or crystal to the other terminal. The terminals are separated by an insulator (usually ceramic) and the intervening space may be filled with a gas, a plastic, or other suitable material. Metal end pieces or caps give mechanical strength and complete the construction of the crystal rectifier.

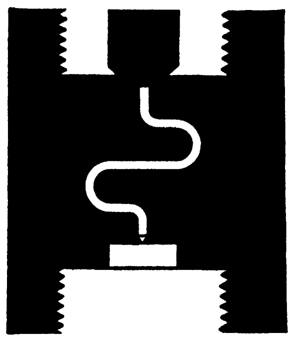


FIG. 1. Typical construction of a crystal rectifier.

When the tungsten terminal is made positive, electrons move in a forward direction from the semiconductor into the tungsten, but, when the polarity is reversed, very few electrons are able to cross the boundary into the semiconductor. The relationship of this unilateral conduction is shown in the curve of Fig. 4 which is typical for crystal and other semiconductor-to-metal rectifiers. This phenomenon of unilateral conductivity is not readily explainable.* The author offers the following. First, it should be noted that the contact between the metal and the conductor is a mere pressure contact, a simple type of boundary, and not a molecular bond. Second, it should be remembered that the contact is a point contact wherein the potential gradient surrounding the point will be high when differences of potential are applied. The electron energy levels which exist in the metal and semiconductor when separated are shown in part *a* of Fig. 2. Here the work function of the metal ϕ_t is relatively low compared to the total work function ϕ_{sc} for the semiconductor. Bringing the metal and semiconductor together in a mere physical contact without any applied potential will not produce much change in the energy levels. However, if a positive potential is applied to the metal, the high potential gradient close to its point will add to the normal electron energy level in the semiconductor near the contact and give a distribution as shown in part *b* of Fig. 2. Under this condition electrons can pass "over the hump" into the tungsten and this constitutes the forward current. When

* For one explanation see "Germanium Crystal Diodes," by E. C. Cornelius, *Electronics*, February 1946.

the polarity of the applied voltage is reversed, the negative potential (gradient) at the tungsten point lowers the energy level of the electrons in the semiconductor as shown in part *c* of Fig. 2, and very few electrons are able to pass to the semiconductor. The latter situation is somewhat analogous to an attempt to remove electrons from a cold electrode.

The crystal rectifier is small in size and light in weight. For use at ultra-high frequencies it is superior to the vacuum diode as a detector because (1) it has smaller input and output capacitances, (2) it does not introduce transit time effects, (3) noise voltages and interelectrode reactances have been reduced, and (4) it does not require any power for heating a cathode.

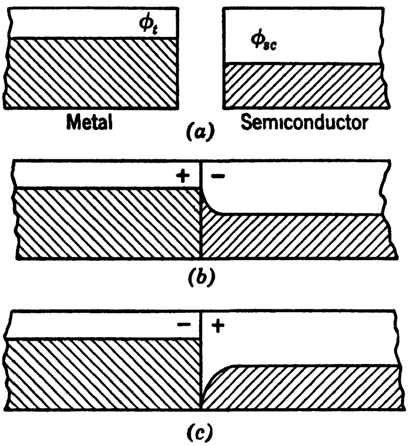
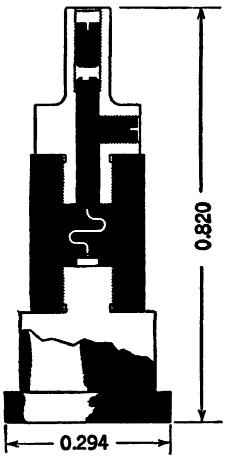


FIG. 2

One type of *silicon* crystal rectifier is illustrated in Fig. 3. It is approximately 13⁄16 inch long and has a maximum diameter of 5⁄16 inch. This rectifier has been designed for easy insertion and plugging into cavity resonators and coaxial butterfly circuits. It is usable as a signal pick-up probe in wave guides and cavities. Silicon rectifiers are designed for use in the first detector stages of super-high-frequency superheterodyne receivers, and for tuned radio-frequency and video detector use at frequencies within the range of 1000 to 25,000 megacycles.

The *germanium* crystal rectifier, or crystal diode, as it is sometimes called, has a cartridge-type construction and the appearance of a one-watt fixed resistor as shown in Fig. 4. A tungsten feeler or catwhisker bears upon a thin disk of optically polished germanium soldered to the end of a wire. The germanium disk is prepared from the natural dioxide form (GeO_2) by reduction in hydrogen, leaving the amorphous metal in a pure state. The dull gray powder is melted and a small amount of tin is added. When it has cooled, a crystallized ingot results which is cut into wafers 0.6 millimeter thick and 3 millimeters square. The polished disks form a lattice-imperfection semiconductor.

The rating and typical characteristics of the selenium crystal diode are given in Fig. 4. This crystal diode has a shunt capacitance of approximately 3 micromicrofarads compared to approximately 15 micromicrofarads for a 6H6 vacuum diode under analogous conditions. This germanium crystal unit is recommended for use between 0 and 100 megacycles though it has been used up to 500 megacycles. It was designed for superheterodyne second detector applications, but it is use-



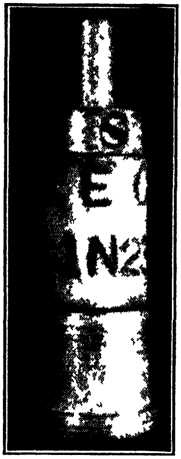


FIG. 3. Typical silicon crystal rectifier. (Courtesy Bell Telephone Laboratories.)

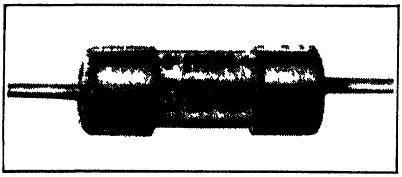
ful in frequency modulation and in television reception where load resistances are low.

Metallic Rectifier Cell.* Several combinations of substances have been discovered which when placed in the form of a sandwich show the property of unilateral conductivity. These sandwich combinations usually consist of a metal, a blocking layer, and some substance which is a semiconductor assembled as shown in Fig. 5. The blocking layer or rectifying layer exists at the surface junction of the metal and semiconductor and is formed by an electrochemical action or chemical action depending on the process for producing the sandwich. Some form of metal electrode is placed in contact with the semiconductor to carry away the current.

This form of rectifying unit conducts electrons fairly readily from the metal to the semiconductor but offers a high resistance to electron movement in the reverse direction. Actually the unit does conduct a

* Also known as barrier cell and as a blocking-layer cell.

little current in the reverse direction and the ratio of the current in forward or conducting direction to the reverse current is called the rectification ratio. It is easy to remember the direction of the electron movement by likening the metal electrode with its plentiful supply of free electrons to the space just outside the thermionic cathode



RATING

Peak inverse anode voltage, max	50 volts
Peak anode current (sine wave), max	60 ma
Average anode current, max	22.5 ma
Surge current (transient peak), max	100 ma
Back conduction at 50 volts, max	2 ma

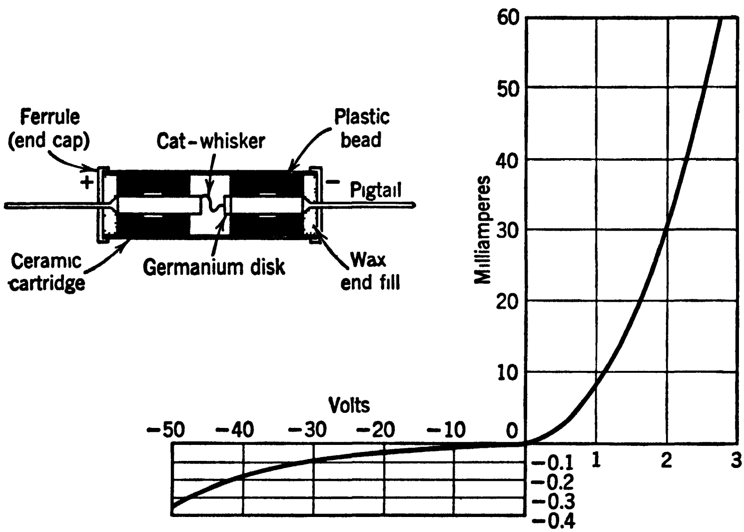


FIG. 4. Construction and characteristics of a typical germanium crystal diode. (Courtesy Sylvania Electric Products, Inc.)

and the semiconductor to the positive anode with a deficiency of available electrons. The theory of the action taking place in the metallic rectifier cell is not fully understood. Curiously, the movement of electrons from metal to semiconductor is opposite to that in the crystal rectifier discussed on the preceding pages. In considering this strange phenomenon, the physical differences in the two kinds of rectifiers should be noted. Thus, (1) the crystal rectifier has a small area of contact, whereas the metallic unit has a large one; (2) the crystal unit has a simple boundary or physical contact, whereas the metallic recti-

fier has a chemical union where the metal and semiconductor have a merging of molecules; and (3) the crystal rectifier has a thick semiconductor unit, whereas the metallic rectifier has a very thin layer. Apparently a polarized film or layer is built up between the metal and the semiconductor by the chemical or electrolytic action which takes place when the sandwich is formed. This layer permits part of the plentiful supply of free electrons in the metal to move across the thin semiconductor when it is positive, but it permits only a few to move under a reversed electric field.

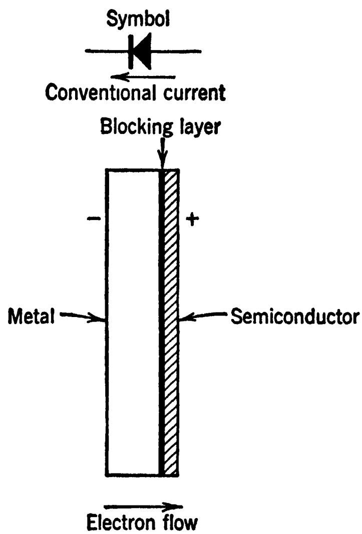


FIG. 5. Elements of a blocking-layer metallic rectifier and its symbol.

The *copper oxide rectifier* is a sandwich consisting of a layer of cuprous oxide on copper. This type of metallic rectifier was invented by Dr. Grondahl and its development followed a discovery made in 1920 while he was studying the photoelectric properties of cuprous oxide. Commercial forms of the copper oxide rectifier cells are shown in Fig. 6. These cells are produced by heating a copper disk or plate in a furnace to a temperature of approximately 1000 degrees F and then quenching it in water. This treatment produces a thin film of cuprous oxide with an outer layer of cupric oxide. The cupric oxide is then removed, leaving a thin layer (about 3 mils or 1000 molecules thick) of cuprous oxide on one side. Contact with the copper oxide surface can be made by holding a lead disk against the oxide surface under pressure or by electroplating a nickel coating on the surface of the oxide. The plated surface is preferred because there is a tendency for the lead disk to "flow" with time and to introduce an undesirable contact resistance in the circuit.

The unilateral characteristic of this cell is shown in Fig. 7 which gives the amperes in the forward direction of current and the milliamperes for the reverse or leakage current. The resistance of the unit is shown in Fig. 8 for a given temperature under a variation of voltage from negative to positive. The resistance also varies with the current

in the forward direction, being approximately inversely proportional to the current, as illustrated in Fig. 9. The resistance of the unit has a negative coefficient of temperature which gives a variation of the

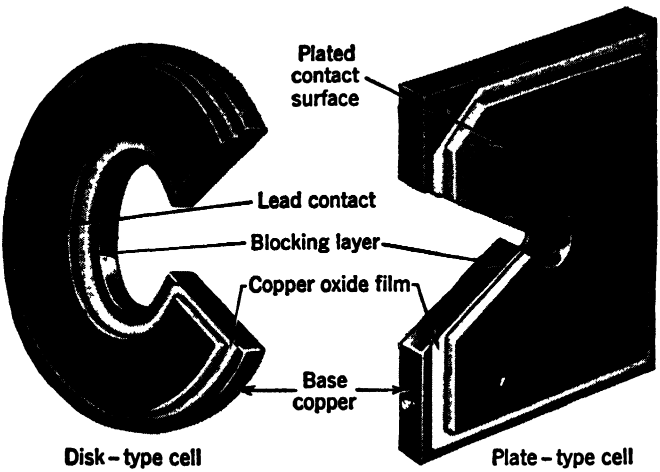


FIG. 6. Cross section of a copper oxide rectifier disk and plate-type cells. (Courtesy General Electric Company.)

current flow in the forward and reverse direction as shown in Fig. 10.

This rectifier is essentially a resistance device. Accordingly, the current squared times the resistance gives the power loss for both

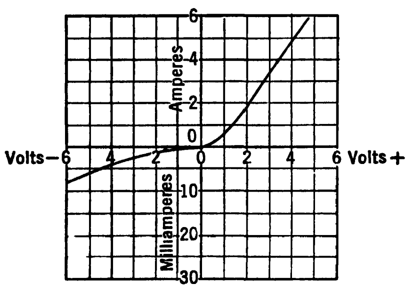


FIG. 7. Current-voltage characteristic curve of a copper oxide rectifier element. (Courtesy General Electric Company.)

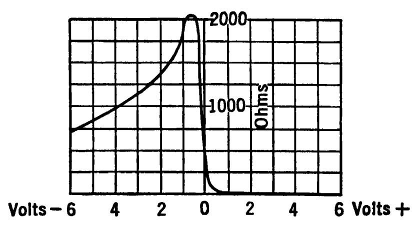


FIG. 8. Resistance-voltage characteristic curve of a copper oxide rectifier element. (Courtesy General Electric Company.)

forward and reverse current flow. These power losses raise the temperature and limit the rectified output of the device.

The various voltage, current, resistance, and temperature characteristics depicted in the figures for the copper oxide unit suggest certain

limitations which must be placed on the design and use of the copper oxide rectifier. In order to prevent an excess of reverse current, which represents a power loss and reduction in efficiency, the voltage impressed

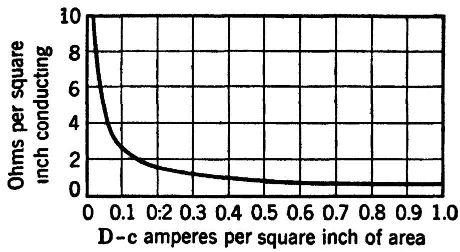


FIG. 9. Resistance-current characteristic of a copper oxide cell at 25 degrees C. (Courtesy General Electric Company.)

across a single cell must be restricted to from 5 to 7.5 volts. Also, in order to limit the leakage current and to prevent ultimate overheating, the maximum temperature for operation of the unit should be limited to about 45 degrees C. The output of the copper oxide rectifier may be increased by improved cooling which can be brought about by the addition of cooling fins and still more effectively by forced air cooling.

The resistance of the copper oxide element increases with age. This "aging" effect on resistance increases with use and with rise in temperature. Fortunately, it appears that the resistance increases to a maximum stable value which is about twice the initial resistance. The result of aging is to reduce the efficiency of the rectifier with time and to change the value of the rectified voltage under load (regulation). Aging seems to be due to some inherent chemical or physical change in the rectifier element for which no remedy has been discovered to date. Commercial rectifier units carry a rating in which the effects of aging have been discounted so that the operation when the unit is new will be much better than the rating and guarantees.

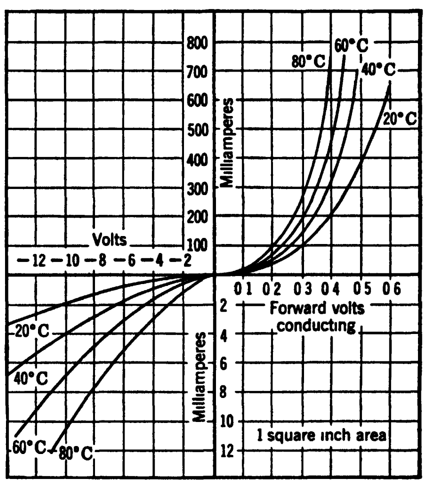


FIG. 10. Current-voltage characteristics of copper oxide rectifier cell. (Courtesy General Electric Company.)

Commercial copper oxide rectifiers usually consist of a number of elements in series or parallel combination. The number of elements in series depends on the applied a-c voltage and the number of elements placed in parallel depends on the required output current. Some copper oxide rectifier assemblies are shown in Fig. 11.

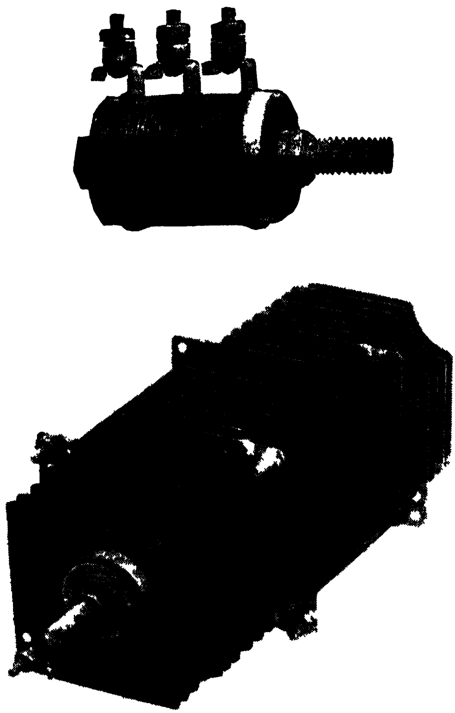


FIG. 11. Copper oxide rectifiers. (Courtesy General Electric Company.)

Copper oxide rectifiers are used in the electric power field for charging batteries, electroplating, motion-picture arcs, corrosion protection, ignitron exciter circuits, and similar applications. Their unidirectional conductivity makes them useful in series with d-c relays to guard against reversed polarity and in many circuits to permit current flow in one direction only. Copper oxide rectifiers are used also to permit measurement of alternating current with a d-c meter element. For this purpose four small copper oxide disks about $\frac{1}{4}$ inch in diameter are connected in the bridge circuit of Fig. 18, center, with a d-c meter in place of load. A similar grouping of four copper oxide disks sealed in a small metal container is known as a "varistor." Varistors are used in

large quantities for modulation and demodulation on long distance telephone carrier circuits.

The *selenium rectifier* is a blocking-layer type of rectifier which uses selenium for the semiconductor and a sprayed metal for a counter electrode. The principle of this device was discovered in the year 1883 though commercial application was not made in Europe until 1930 and in America a few years later. The selenium rectifying cell is illustrated in Fig. 12. It consists of a back plate of either iron or aluminum covered by a thin film of selenium. The film is given a series of controlled heat treatments to reduce the selenium to a suitable crystalline form. Next, a low melting alloy is sprayed onto the selenium surface. This layer is called the counter electrode and constitutes the active metal electrode of the sandwich unit. Subsequent electrochemical processes produced by the application of alternating current to the cell form a film or blocking layer between the counter electrode and the selenium.

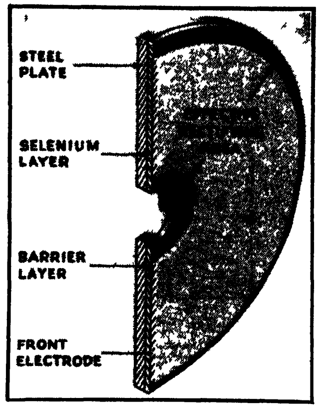


FIG. 12. Selenium rectifier cell (Courtesy International Telephone and Telegraph Company)

The current-voltage characteristic of a selenium cell per square inch of contact surface is shown in Fig. 13. The nature of the unilateral conduction is the same as that of a copper oxide unit with the electrons moving from the counter electrode into the selenium. The selenium cell has the same general characteristics as the copper oxide unit such as negative temperature coefficient of resistance, inverse relationship between resistance and current, and increase of resistance of the unit with age. However, it does differ from the copper oxide unit in some respects. It will withstand a maximum peak voltage of 25 volts per cell without damage. This means that for the same voltage output only one-half to one-third as many cells need be placed in series. Second, the selenium cell may be operated safely up to a maximum temperature of 75 degrees C. However, the continuous current output must be reduced as the temperature rises, so the advantage of higher temperature is not so important. The weight of the selenium rectifier for a given rating may be only one-seventh that for copper oxide for steel back plates or one-half this ratio if aluminum back plates are

used. One disadvantage of the selenium unit is that the resistance of the blocking layer is lower for direct current than for alternating current and that the blocking-layer resistance deteriorates when the cell is not in use though it is restored in a few seconds with the application of an a-c voltage. Hence the cell is not suitable for operating relays for reversed polarity on d-c circuits. Otherwise, the applications of the selenium rectifier are the same as those for the copper oxide type. The parts and assembly of a selenium rectifier unit are shown in Fig. 14.

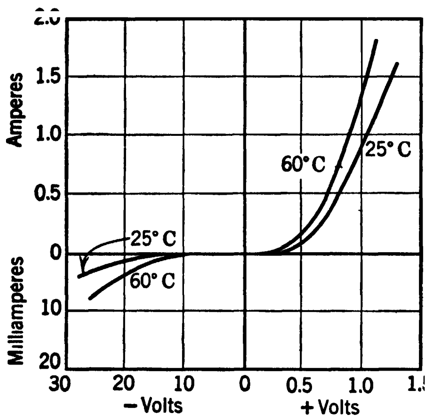


FIG. 13. D-c voltage characteristic of a selenium rectifier cell. (Courtesy General Electric Company.)

In 1946 new miniature selenium rectifier units were placed on the market which perform the same function as the 35 and 117 series of rectifier tubes. These new selenium units are more rugged and have a longer life than the rectifier tubes which they are designed to replace. One of these new units is illustrated in Fig. 15.

The *magnesium-copper sulphide rectifier* was invented in Austria and is manufactured in the United States under basic patents issued to Samuel Ruben in 1925. This blocking-layer metallic rectifier uses a magnesium plate for the metallic electrode and copper sulphide for the semiconductor. A back plate or electrode is placed in contact with the copper sulphide for collecting the current. The three pieces are assembled on a stud with radiator plates and held under pressure as illustrated in Fig. 16. The active blocking layer is formed between the magnesium and the copper sulphide by an electrochemical process. The forward movement of the electron flow is from the metal magnesium to the semiconductor copper sulphide. This rectifier unit must be protected from absorption of moisture to insure a long life; this necessity constituted one of the problems in the early manufacture and use of this rectifier. This rectifier cell has a rectification ratio of 75 to 1, which is lower than that of other disk rectifiers and tends toward a lower rectifier efficiency.

The advantages claimed for the magnesium-copper sulphide rectifier are (1) its ability to withstand heavy current overloads, (2) an

extreme range of operating temperature, plus (3) a high normal current density. This unit has a self-healing rectifying film which will

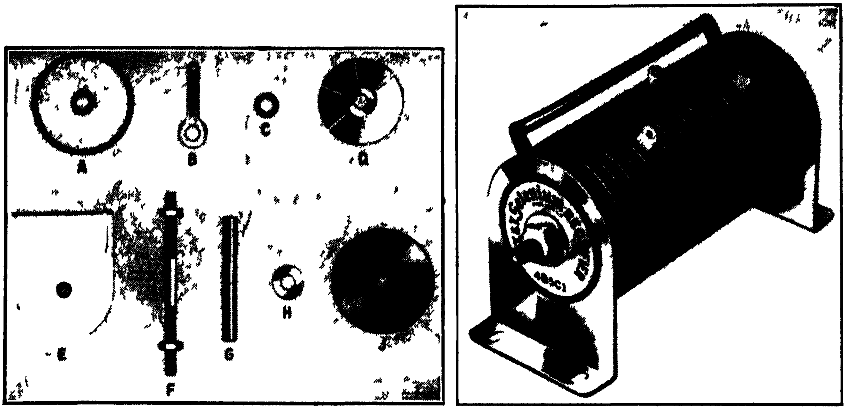


FIG 14 Parts of a selenium rectifier and an assembled selenium rectifier (Courtesy International Telephone and Telegraph Company)

seal a puncture due to overload or a transient voltage surge. This rectifier will operate under a range of temperature from -70 to 135 C and has been tested in the laboratory under loads that produced temperatures as high as 275 degrees C. The normal current density depends on the method of cooling employed and lies in the range of 25 to 50 amperes per square inch.

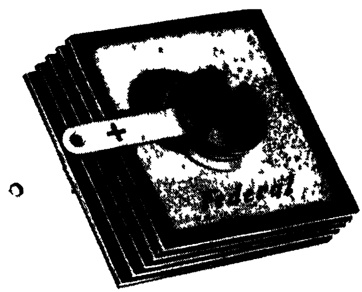


FIG. 15. Miniature selenium rectifier (Courtesy International Telephone and Telegraph Company.)

The magnesium-copper sulphide rectifier may be employed for the same applications as the metallic rectifiers previously discussed. A commercial rectifier of this type is shown in Fig 17.

Summary on Metallic Rectifiers. Metallic rectifier cells are resistive units, and hence provide a unity power factor load as far as the direct rectification process is concerned. These cells can be used in series and parallel combination for any type of rectifier circuit. Three typical circuits for metallic rectifiers are shown in Fig. 18. Several other circuits will be given in Chapter XII.

The design of a metallic rectifier for a given application will be determined by the load amperes, the desired voltage regulation, and the

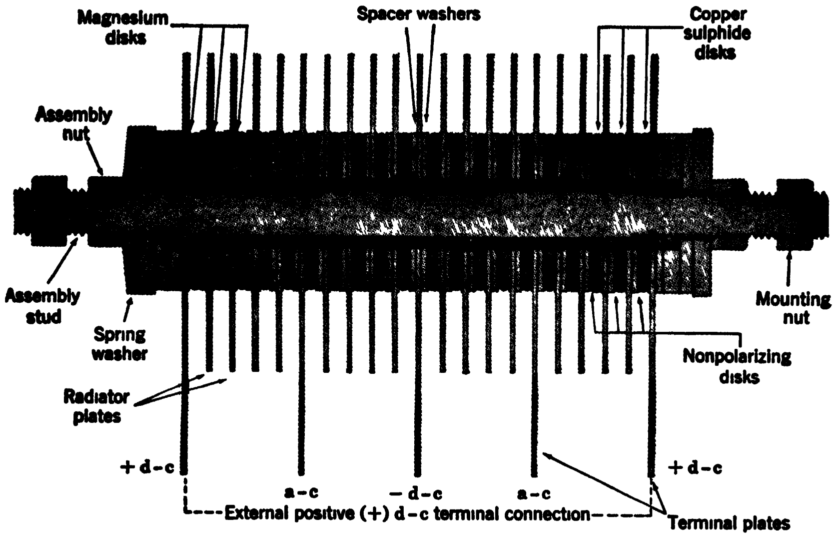


FIG. 16 Construction of a magnesium-copper sulphide rectifier (Courtesy P R Mallory & Company, Inc)

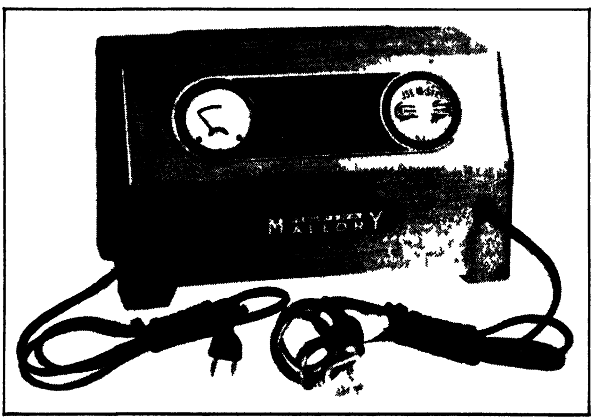


FIG. 17 Commercial battery-charging unit using magnesium-copper sulphide rectifying disks. (Courtesy P R Mallory & Company, Inc)

characteristics of the type of rectifier cell employed. The important cell characteristics are (1) the permissible inverse peak voltage, (2) the permissible temperature rise, and (3) the permissible current density. All metallic rectifier cells have negative coefficients of resistance

though the resistance is approximately constant in the normal current density range. If the resistance is considered constant, the voltage regulation curve would show a linear fall of potential with load. The negative temperature coefficient of resistance (Fig. 9) causes the regulation to be better than linear, as shown in Fig. 19. To provide satisfactory regulation the total resistance of the rectifier cells at full d-c

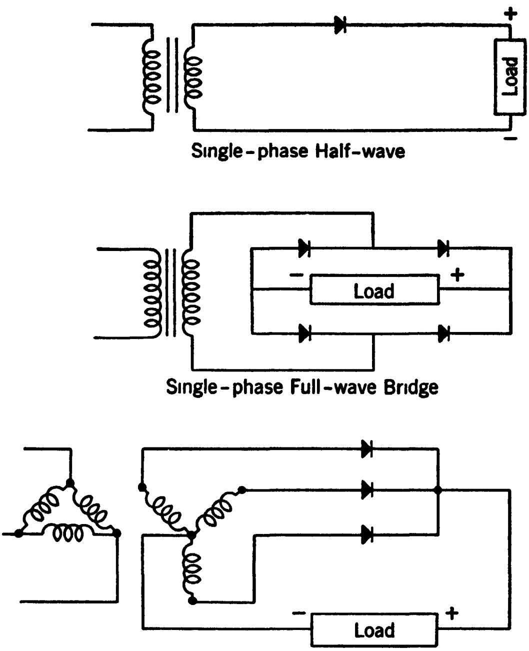


FIG. 18 Typical circuits for metallic rectifiers.

load should not exceed 10 to 15 per cent of the load circuit resistance. The number of rectifier cells required is determined by the permissible inverse peak voltage and the d-c no-load voltage. Lastly, the size (and number in parallel) of the rectifier cells is determined by the d-c load current. The manufacturers use empirical methods for designing rectifiers for applications. The student may apply the preceding statements of this paragraph, the data on Figs. 9, 10, and 13, plus the data of Table 1 to make an approximate design or to compute the probable rating of a rectifier having an unknown rating.

The efficiency of a metallic rectifier is the ratio in per cent of the d-c power output to the total power input where both output and input are measured with a-c wattmeters. The efficiency of metallic rectifiers

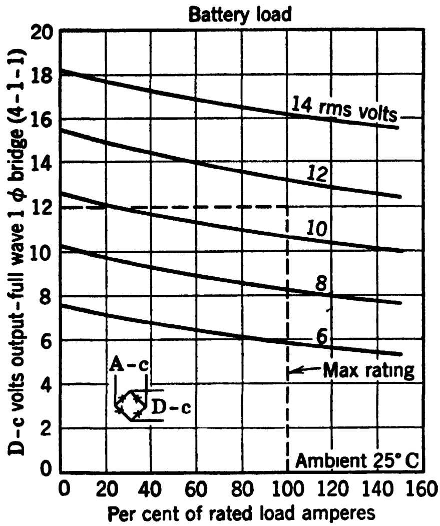


FIG. 19. Voltage-regulation curves of a selenium rectifier. (Courtesy General Electric Company.)

is reduced by the reverse current flow as well as by the I^2R loss within the rectifier units. Efficiency of rectifiers will be given further consideration in Chapter XII. Since the term "efficiency" as applied to metallic rectifiers is subject to different interpretations, the manufacturers of metallic rectifiers prefer to use the term "conversion ratio." *The conversion ratio of a metallic rectifier is the ratio, in per cent, of the product of the average values of d-c voltage and d-c current output to the total a-c power input.*

A summary and comparison of the characteristics of copper oxide, selenium, and magnesium-copper sulphide rectifiers is given in Table 1. The student should study this table.

TABLE 1.* COMPARISON OF METALLIC RECTIFIERS

Type of Metallic Rectifier Cell	Peak Inverse Volts per Cell		Current Density Amperes per Square Inch (full-wave rectification)		Maximum Recommended Temperature (degrees)		Efficiency (overall per cent)	
	Normal operation	Intermittent service or reduced life	Convectional cooling	Forced draft cooling	Normal operation	Intermittent service or reduced life	For 1 ϕ	For 3 ϕ
Copper oxide	8.75–11.5	30.0	0.3	0.5	35 C	45 C	60–70	75–80
Selenium	25.0	54 0	0 3	0.75	35 C	75 C	60–70	70–75
Magnesium-copper sulphide	5.0	5 0	25.0	50 0	40 C	85–130 C	37–45	52.7

*** Several factors enter into the rating and operation of metallic rectifiers so that the values suggested in this table should be considered as approximate.**

Problems

In the solution of the following problems, assume that the peak applied voltage is 1.41 times the applied rms alternating voltage and that the d-c average voltage is 0.318 times the peak for half-wave rectification and 0.628 times peak for full-wave.

1. An a-c voltage of 60 volts (rms) is to be rectified by copper oxide cells to deliver 2.0 amperes of direct current with full-wave rectification under normal operation and with convectional cooling. Calculate (a) the number of cells, (b) the area of the cells, and (c) the average d-c voltage under load.

2. An a-c supply of 120 volts rms is applied to selenium rectifier stacks (full-wave) to give 100 milliamperes direct current. How many cells will be required and what should be their size for intermittent service? What will be the d-c average voltage? What should be the magnitude of the load resistance? What will be the resistance drop across the rectifier?

3. A metallic rectifier is to be used near a furnace where the ambient temperature may rise to 100 degrees C. What type of cell should be used?

4. Using Fig. 19, calculate the per cent voltage regulation of a selenium rectifier which charges a 6-volt storage battery (with taper charge) up to a near zero rate with a final voltage of 7.5 volts.

5. Repeat Problem 4 for a 12-volt battery, having a final (zero charging current) voltage of 15.5 volts.

REFERENCES

Hamann, C. E., and E. A. Harty, "Fundamental Characteristics and Applications of Copper-Oxide Rectifiers," *Gen. Elec. Rev.*, August 1933.

Clark, Carole A., "Selenium Rectifier Characteristics, Applications, and Design Factors," *Elec. Commun.*, **20**, 1941.

Cornelius, E. C., "Germanium Crystal Diodes," *Electronics,* February **1946**, p. 118.

Chapter IX

ALTERNATING-CURRENT THEORY

The performance of electronic equipment is determined by the characteristics of the associated circuits and devices as well as by the characteristics of electron tubes. In like manner, the performance of other types of electric control equipment is governed by the characteristics of the associated circuits and components. Hence any understanding of industrial electronics and control equipment requires a knowledge of a-c circuit theory and the theory of many components used in such equipment. The reader is presumed to have a general knowledge of a-c theory so that only a short summary of a-c circuits plus a discussion of some phenomena of special importance in control equipment will be reviewed at this point.

A-C Theory. An a-c circuit may consist of resistors, inductors, capacitors, or combinations of these three units. The characteristics of these units are reviewed in Fig. 1. For a noninductive resistance the current-time change is in phase with the voltage; in a pure inductance the current lags the voltage by 90 electrical degrees; and in a capacitance the current leads the applied voltage by 90 electrical degrees.

Circuits and circuit elements may be classified as *linear* and *nonlinear*. Under common usage a circuit or element is considered to be linear when the voltage-current relation follows Ohm's law ($I = E/Z$). Thus a pure resistance, a pure inductance, and a capacitance are generally considered to be linear elements in an a-c circuit. Similarly, electron devices that have unilateral conductivity, changes in amplifying power, changes in mutual conductance, changes in plate resistance, and so forth, are usually classed as nonlinear elements.

Under operating conditions the magnitude of circuit elements often becomes a function of temperature, frequency, flux density, and other factors. Accordingly, an analysis of the properties of circuit elements under varying conditions becomes necessary in the study of control. The resistance of pure metals rises with temperature and for the usual operating range may be considered as a straight-line function. Alloys

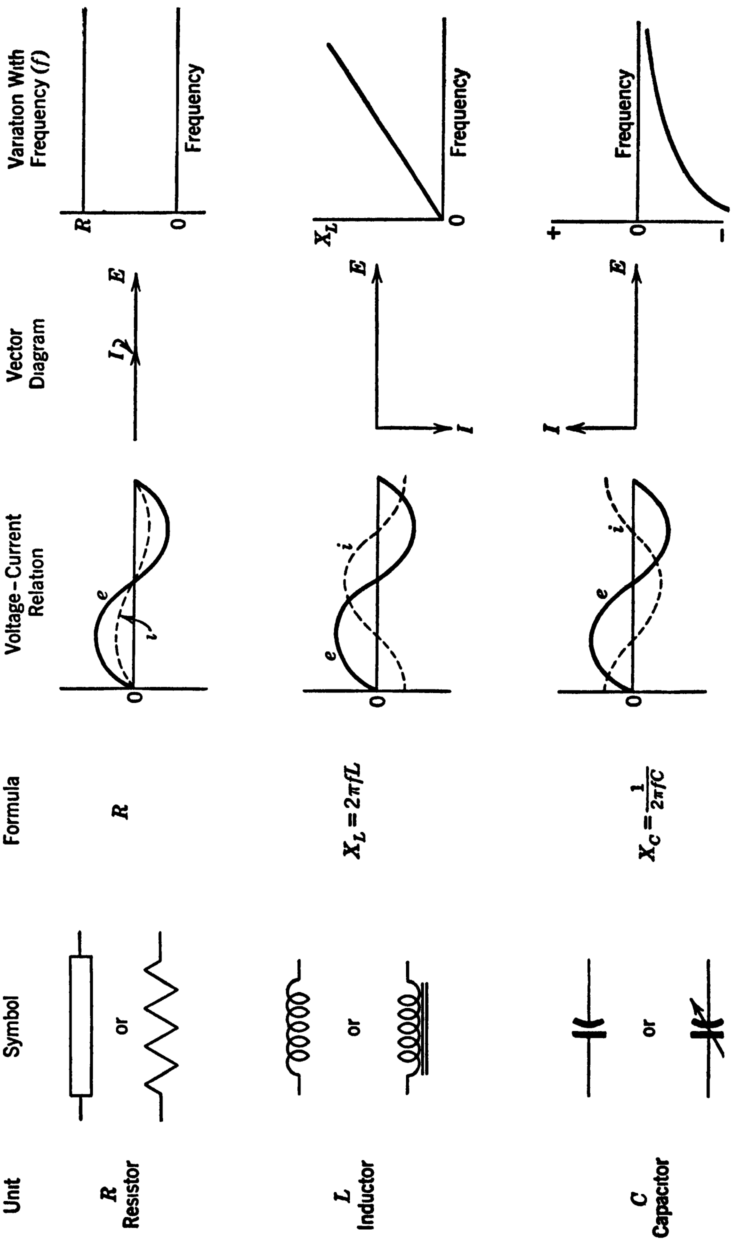

Unit
Symbol
Formula
Voltage-Current Relation
Vector Diagram
Variation With Frequency (f)
R Resistor
or
R
e
i
0
I
E
R
0
Frequency
L Inductor
or
$X_L = 2\pi fL$
e
i
0
I
E
X_L
0
Frequency
C Capacitor
or
$X_C = \frac{1}{2\pi fC}$
e
i
0
I
E
+
0
−
Frequency

FIG. 1

of metals and some combinations of materials have variations of resistance with temperature which make them useful in control circuits to be considered subsequently. For the low frequencies used in power circuits the magnitude of resistances may be safely considered as constant, as shown in the upper right corner of Fig. 1. However, for the higher frequencies to be found in electronic circuits the ohmic resistance of resistors rises with frequency due to the so-called skin effect. This effect is illustrated in the five views of Fig. 2. In part 1 of Fig. 2, it is assumed that an instantaneous current of uniform density is flowing throughout the cross section of the conductor. This current

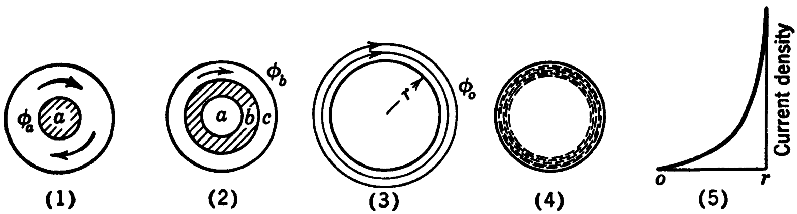


Fig. 2. Skin effect for a copper conductor.

will produce a magnetic field around the elemental cylinder a as shown, and the changing flux will induce an emf which will oppose the flow of current in a. In part 2, the elemental tube b carries current which causes flux ϕ_b to surround itself. The outer tubular element c is surrounded by the external flux ϕ_o shown in part 3. Cylinder a is surrounded by the flux produced by itself plus that produced by the current flowing in the tubes b and c. Hence in tube a the flow of current is opposed by the maximum or total emf of self-inductance due to the total surrounding flux. The current in b is opposed by the flux surrounding it. The effect of this self-inductance within the conductor is to distort the current density within the conductor, crowding the current toward the outside of the conductor as shown in parts 4 and 5 of Fig. 2. For very high frequencies nearly all the current flows on the surface so that a thin copper tube may have nearly the same resistance as a solid copper conductor of the same diameter. A thin coating of graphite or other conducting material on the surface of a cylinder of insulating material will make a suitable resistor or conductor for high frequencies. Thus the ohmic resistance of conductors may vary nonlinearly with the frequency of the applied voltage. The magnitude of an inductive reactance having a pure inductance $(2\pi fL)$ varies along a straight line with frequency, as depicted in Fig. 1.

It is impossible to construct an inductance or choke coil that does not have some ohmic resistance, because the coil must consist of turns of a conductor and all metallic conductors have some resistance. Thus all inductance coils may be resolved into components of a resistance R and an inductance L in series. For most applications in electronics it is desirable to keep the resistance of choke coils as low as possible and to keep the inductance L as high as possible. Thus the ratio of L/R is important in coil design. The ratio $2\pi fL/R$ is called the figure of merit, or Q, of a coil. In general, it is desirable to use coils having a Q of 100 or higher. For industrial applications the Q of a coil is often expressed as the volt-amperes divided by the watts or VI/W. This is an approximate relationship that is derived from the first ratio as follows:

$$Q = \frac{2\pi fL}{R} = \frac{2\pi fLI \times I}{R \times I^2} = \frac{\text{volt-amperes}}{\text{watts}}$$

Inductance coils are frequently used in series and parallel resonant circuits where the loss in the capacitor is close to zero and the Q of the coil becomes the VI/W for the resonant circuit.

The inductance L of a choke coil having an iron core varies with the degree of saturation of the core. A typical magnetization curve for iron is shown in Fig. 3. A choke coil having an iron core and operated at a constant frequency will have a constant inductance as long as the flux in the core operates over the range from a to b on Fig. 3, since the magnetizing field H and the resulting flux will be proportional to current in the coil. When the iron core is operated over the bend of the magnetization curve from b to c the inductance will be changing rapidly (nonlinear). After the iron becomes saturated (part c on curve), further increase in magnetizing field H will produce the same flux change as would exist in air. For this region the value of incremental L again becomes constant or linear. The change in L due to the saturation of the iron is applied in the saturable core reactor to be explained subsequently. With the rise in frequency another form of skin effect takes place in the iron core which forces the

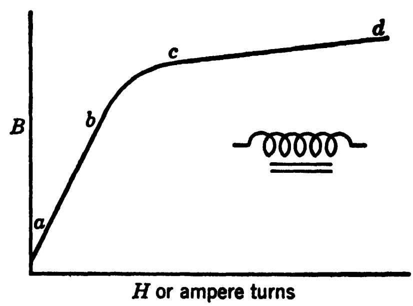


FIG. 3

flux to the outside and makes the iron core less effective. The theory of this form of skin effect will be explained under induction heating. Ordinary iron cores cannot be used effectively in coils for frequencies above 10,000 cycles. However, powdered-iron cores may be used for frequencies up to 100 megacycles per second.

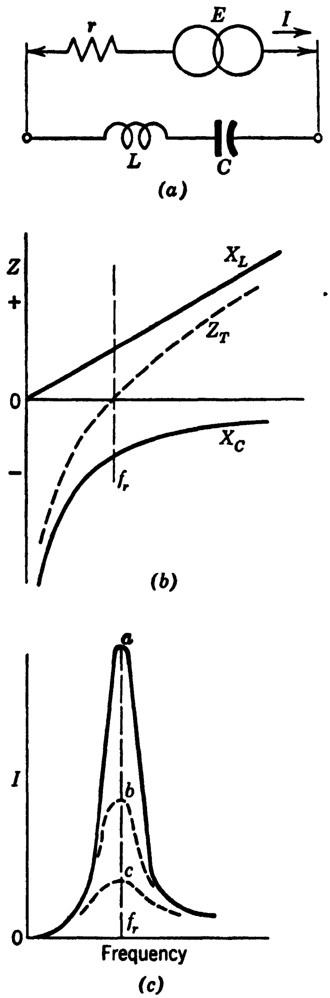


FIG. 4. Series resonance.

The theory of skin effect explained under the preceding discussion of resistance variation also applies to the self-inductance of a conductor itself. Self-inductance is due to the flux linking with the current in the conductor $L = N(d\phi/dI)$. The changing flux may exist both within and without the conductor and the magnitude of L is due to both the internal and external linkages. The external flux for a given current (where $\mu = 1$ for external medium) is independent of the current distribution within the inductor but the internal flux will depend on the current distribution. Thus if a high frequency results in current conduction on the skin or surface of the conductor, the self-inductance will decrease nonlinearly with the rise in frequency.

The voltage drop across a capacitance for a fixed frequency is directly proportional to the current flow. However, the reactance of a capacitance is inversely proportional to the frequency and is a nonlinear function as shown in Fig. 1. This nonlinear function becomes very useful in resonant circuits.

Resonant circuits consist of combinations of inductance L and capacitance C plus some resistance R since inductances always contain some ohmic resistance. Typical characteristics of the series resonant circuit are shown in Fig. 4. Assume that the voltage across the LC combination is held constant and the frequency is varied from zero. The reactances of the inductance X_L and capacitance X_C will vary as

previously shown in Fig. 1 and repeated in center view of Fig. 4. The impedance of the series combination Z_T equals $X_L + (-X_C)$ and is depicted by the dotted line. For one value of frequency f_r the inductive reactance X_L equals the capacitive reactance X_C, giving zero impedance for the circuit. This represents the condition of series resonance, and the frequency f_r is called the resonant frequency. The current I flowing in the series circuit for the varying frequency is shown in the lower part of Fig. 4. When the frequency is zero the capacitance blocks the current (direct) because X_C is infinite. As the frequency rises C conducts and as the frequency of resonance is approached the current rises rapidly, reaching a peak at the resonant frequency (if L and C had zero resistance the current would rise to infinity provided the impressed voltage were maintained). Beyond resonant frequency the current falls off rapidly. Since the circuit always has some small resistance the resonant peak rises to some point such as a. If the resistance R of the circuit is increased the resonant peak will be reduced to points such as b and c and the width of the peak will be increased. Obviously, the series resonant circuit is highly selective for a narrow band of frequencies and is used where a selective filter is desired.

The resonant frequency depends on the magnitude of L and C and can be made to vary over a wide range by making either L or C variable. Resonance exists when

$$X_L = X_C$$

$$2\pi f_r L = \frac{1}{2\pi f_r C}$$

$$f_r = \frac{1}{2\pi\sqrt{LC}}$$

where f_r is the frequency of resonance.

A series LC circuit without resistance and carrying a current I is shown in Fig. 5. At the resonant frequency the voltage drops across the elements will be

$$V_L = IX_L$$

$$V_C = IX_C$$

$$V_{ab} = V_L + V_C$$

and at resonance

$$X_L = -X_C$$

$$\therefore V_{ab} = 0$$

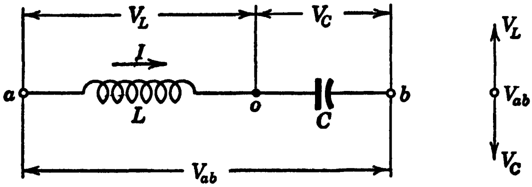


FIG. 5

Since there is no difference of potential between points a and b, these points may be connected together as shown in the upper part of Fig. 6. The new connection gives a parallel LC circuit having a current circulating within itself. There will be a voltage drop across the parallel circuit equal to $IX_L = -IX_C$, but no current will flow outside the parallel branches. The impedance Z of the parallel circuit by Ohm's law is

$$Z = \frac{V}{I} = \frac{IX_L}{0} = \infty$$

or the impedance Z of the parallel circuit may be calculated thus:

$$Z = \frac{X_L X_C}{X_L + X_C}$$

At resonance, $X_L = -X_C$; hence

$$Z = \frac{-X_L^2}{0} = \infty$$

Thus a parallel circuit consisting of pure L and C will offer an infinite impedance at the frequency of resonance. Since the series resonant circuit offers zero impedance at resonance it is the usual practice to speak of parallel resonance with an infinite impedance as *antiresonance.* The current in the parallel branches of L and C and the line current for a variation of frequency is shown in part b of Fig. 6. The variation of the impedance of the parallel LC circuit with frequency has the peaked value shown in part c of Fig. 6. The presence of resistance in the parallel circuit causes some power loss and serves to reduce the value of the circulating current and the impedance across the circuit at resonance. The parallel LC circuit is frequently called a *tank circuit.* It is often used as a filter for an undesired band of frequencies.

The critical or resonant frequency for the antiresonant parallel or tank circuit is calculated by the same formula (1) as for the series re-

sonant circuit if the resistance is negligible. The antiresonant condition can be produced in an LC parallel circuit for a wide range of frequencies by a variation of the magnitude of L or C as in the series LC circuit.

A mechanical analogy to the tank circuit is the 4-cycle gasoline engine. During the combustion stroke of the piston, kinetic energy is stored in the rotating flywheel. For the next three strokes (exhaust, intake, and compression) the flywheel must supply all the losses of the engine and load. As kinetic energy is lost the angular velocity of the flywheel decreases. Hence, the speed between explosions is not constant, and the greater the load for a given flywheel, the more pronounced the variation in speed In the vacuum-tube circuit the function of the flywheel is performed by the tuned "tank" wherein energy is stored alternately in the capacitor and inductance, causing the current to oscillate, flowing on one-half of the cycle to the condenser and on the next to the inductance. It should be noted that the tank current is a true alternating current and generally is many times the magnitude of the current feeding into the tank circuit.

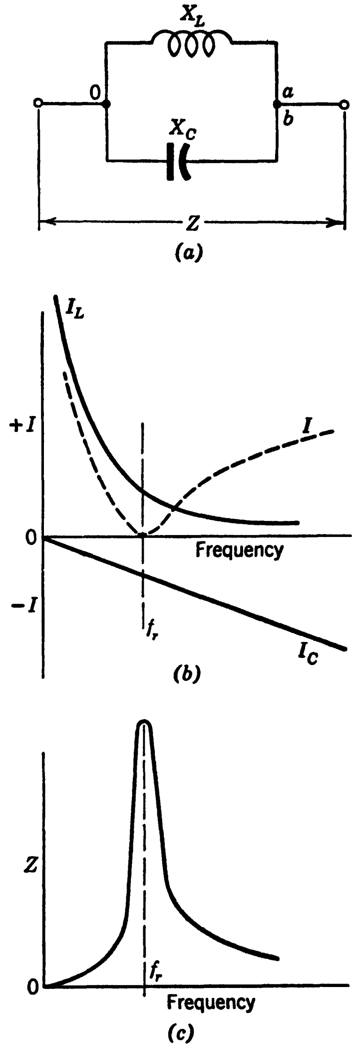


Fig. 6. Antiresonance.

Resonant circuits have many applications. They may be used (1) for frequency discrimination (selection or rejection), (2) for impedance matching in connecting two units such as generator and load, and (3) for voltage amplification.

Whenever a series or parallel LC circuit has its value of L or C adjusted to be in resonance for a given frequency, it is said to be *tuned* for that frequency. As such it is called a tuned circuit and *tuned circuits* play a very important part in electronics.

Problems

1. A choke coil wound on an iron-dust core has a resistance of 65 ohms and an inductance of 40 millihenries. Neglecting skin effect, what is the Q of the coil at 60 cycles, 1200 cycles, and 100,000 cycles?

2. A series resonant circuit consists of a resistor (5 ohms), inductor (50 millihenries), and a capacitor (1.5 microfarads). What is the resonant frequency? What current will flow when 115-volt, 60-cycle alternating current is impressed across the circuit?

3. Design a series resonant filter to pass a frequency of 50,000 cycles, using a 1-microfarad capacitor and an inductance L of ? millihenries. What will be the impedance of this circuit at 45,000 cycles?

4. A tank circuit consists of a choke coil of $L = 5$ millihenries and $R = 0.005$ ohm, and a capacitor of 0.4 microfarad. Calculate the resonant frequency. If a circulating current of 100 amperes (effective) flows in the tank circuit, what is the power loss due to this current?

5. What is the impedance of the tank circuit of the preceding problem at a frequency 90 per cent of the resonant value? at a frequency of 125 per cent?

6. A choke having an inductance of 7 millihenries is to be used in a circuit tuned for 5000 cycles. Calculate the other element of the circuit.

Chapter X

COMPONENTS AND CIRCUITS FOR CONTROL

Components and Circuits for Control. The marvelous achievements in the field of industrial electronics and control are made possible by the use of many component circuits and devices. Several of these components may be employed in combination to achieve the desired result. An understanding of the complete equipment will be greatly simplified by a study and classification of the individual component circuits and devices. The complete equipment may be divided and classified in numerous ways. For the purpose of this text the following classification will be made and each of the twenty items listed will be treated in sequence before passing to the important applications in the industrial field.

1. Vacuum-tube amplifiers.
2. Vacuum-tube oscillators.
3. Saturable-core reactors.
4. Series impedance transformers.
5. Peaking transformers.
6. Phase-shift circuits.
7. Free-wheeling circuits.
8. Rectifiers.
9. Filters.
10. Voltage dividers.
11. Time-control circuits.
12. Constant-voltage circuits.
13. Nonlinear resistors.
14. Electronic switches.
15. Long-tailed pair.
16. Light-control devices.
17. Temperature-control devices.
18. Position-control devices.
19. Rotary amplifiers.
20. Antihunt circuits.

1. Vacuum-Tube Amplifiers. The vacuum triode, tetrode, and pentode are amplifying tubes. Associated with suitable circuits they constitute one type of amplifier and as such they have revolutionized the art of electrical communication, measurement, and control. The function of the amplifier is to increase or magnify a very weak signal until it is capable of controlling sufficient electric power for producing light, sound, or mechanical work. An amplifier does not create energy but controls sources of electric energy to give outputs that follow the original input signal with fidelity. An amplifier performs its function in one or in a series of stages or steps. The amount of amplification

that can be attained for some applications is limited by various "noise currents" which may be introduced by the tubes employed in the amplifier. Such noise currents are (1) shot noise arising from random electron movements in conductors or electrodes, (2) microphonics resulting from vibration of tube parts, and (3) hum caused by stray magnetic and electric fields.

The individual stage of an amplifier may be designed for (1) voltage amplification or (2) power amplification. Voltage amplification signifies that the output voltage across the load is greater than the voltage of the input signal and that the power output is small. Power amplification means that the output power (volts times amperes) is large compared to that of the input signal. Frequently a complete amplifier consists of one or more stages of voltage amplification followed by a final stage of power amplification. Any amplifying tube may be provided with a circuit to give either voltage or power amplification. However, the power output of a tube is determined by its size, its construction, and the heat dissipation of its plate. Hence design practice usually employs a small tube of low plate dissipation for voltage amplification and a larger tube of high plate rating for the power amplification stage. The overall power gain in an amplifier of several stages may be very large, approaching infinity, because the input to the first stage of voltage amplification may approach zero, whereas the energy output of the final stage may be large.

The circuits and sometimes the tubes employed in amplifiers depend upon the range of frequencies to be handled. The classification of amplifiers on this basis may be summarized as follows:

Direct-coupled amplifiers	0–15,000 cycles.
Audio-frequency amplifiers	20–15,000 cycles.
Radio-frequency amplifiers	Frequencies above the audio range with each amplifier designed to handle only a narrow band of frequencies.
Video-frequency amplifiers	0–10,000,000 cycles.

Vacuum tubes used in amplifiers are operated with a normal negative grid bias (i.e., with the grid held negative when the input signal voltage is zero). The magnitude of the negative grid bias with respect to the transfer characteristic of the tube determines the amplifying characteristic obtained and also determines the classification of the amplifier stages. The terminologies for amplifiers are Class A, AB, B, and C, each to be discussed in a subsequent section.

The basic circuit for a single stage of an amplifier is given in Fig. 1.

The grid is given a negative bias by the battery E_{cc} and the plate is supplied by the battery E_{bb}.* An a-c input signal voltage e_g is impressed across R_g and thence into the grid-cathode circuit. This varying input voltage will cause a varying current in the plate-cathode circuit, including the load resistance R_L. The current through R_L will produce a varying voltage drop e_z. If the plate current is small and R_L is large, the i_bR_L drop may be several times as large as the input e_g and voltage amplification results. If R_L is a relay or a loudspeaker or other device having a low or moderate resistnance ad if i_b is relatively large, considerable power $i_b{}^2R_L$ will be delivered to R_L and power amplification results.

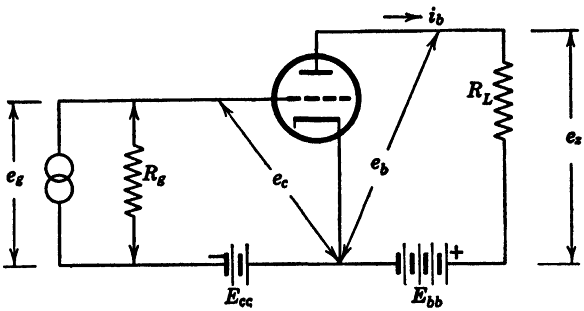


FIG. 1. Basic amplifier circuit.

Thus there is no difference in the appearance of the basic circuit for voltage or power amplification, but there is a difference in the magnitude of the load resistance.

The action taking place within the amplifier stage can be visualized through the picture of phase relationships of voltages and currents shown in Fig. 2. A sine-wave signal (shown about zero axis) is impressed in series with the negative grid bias giving the grid voltage shown at the bottom (e_c). The resulting plate current (a-c component) i_p is in phase with the grid voltage, but the varying plate voltage e_b is opposite in phase (180 degrees) to the plate current and grid voltage. Similarly, the load voltage e_z ($e_z = E_{bb} - i_pR_L$) is opposite in phase to the grid input voltage. *Thus the amplified output voltage is 180 degrees out of phase with the impressed input signal voltage.*

The analysis of the basic amplifier circuit can be simplified by the use of the equivalent circuit shown in Fig. 3. In this equivalent circuit the tube is replaced by a resistor having the same value as the acplate resistance (dynamic) r_p and a generator of voltage μe_g (amplifica-

* See front of book for standard symbols for electron-tube circuits.

tion factor μ times signal voltage) connected to a load of impedance Z_L. If the load is a pure resistance R_L, the plate current will be

$$i_p = \frac{\mu e_g}{r_p + R_L} \tag{1}$$

where e_g and i_p are the instantaneous values of the varying components, and it should be remembered that the current considered is *electron*

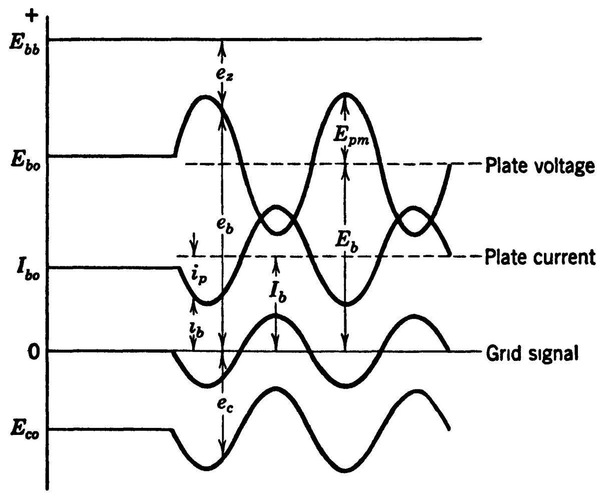


FIG. 2. Amplifier current and voltage relationships.

current flowing from cathode to plate. Such a current, flowing through an impedance, will produce a voltage rise in the direction of current flow. Therefore the varying component of the load voltage across the load resistance will be

$$e_z = -i_p R_L = -\frac{\mu e_g R_L}{r_p + R_L} \tag{2}$$

Dividing both sides by e_g gives the voltage gain G.

$$G = \frac{e_z}{e_g} = -\frac{\mu R_L}{r_p + R_L} \tag{3}$$

The negative sign indicates the 180-degree phase reversal between the grid and plate voltages. Since the equivalent circuit considers only the cathode-to-plate part of the tube, it may be applied to tetrodes and pentodes as well as to triodes.

Amplifier Classifications. A *Class A amplifier* is one in which the grid-bias and alternating grid voltages are of such magnitudes that plate current flows throughout the cycle The Class A operation of a tube is illustrated graphically by the dynamic transfer characteristic of Fig 4 An examination of this graph will show that plate current flows during the positive and negative half-cycles of the a-c signal voltage Since the i_b-e_c curve is not linear over its entire length, the tube must be biased so that the signal operates over the straight-line section in order to give an exact reproduction of the input wave form If the grid of the tube is biased incorrectly so that the grid voltage acts over the nonlinear portion of the curve, a distorted plate-current wave form will result (Fig 5) In this current wave, the positive and negative half-cycles are unequal and the result is known as amplitude distortion If this distorted current is passed through a load resistance, the resulting load voltage will have a like wave form Distortion will occur also in a Class A amplifier if too large a value of a-c signal voltage is applied to the grid of the tube so that operation takes place over both the linear and nonlinear portions of the curve The maximum power output that can be obtained from any amplifier stage will depend on the efficiency of the circuit and the permissible plate-dissipation rating of the tube The plate efficiency of an amplifier stage is the ratio of the a-c power output (at signal frequency) available at the load to the total plate power input (d-c plate voltage times plate current) expressed in per cent The plate dissipation of a tube is the difference between the power input and the power output. The efficiency of Class A amplifier stages is of the order of 20 to 25 per cent Since the Class A amplifier gives little distortion when properly operated, it is widely used for speech amplification and voltage amplification at radio frequencies.

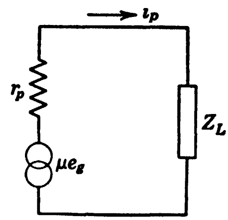


FIG 3 Equivalent circuit

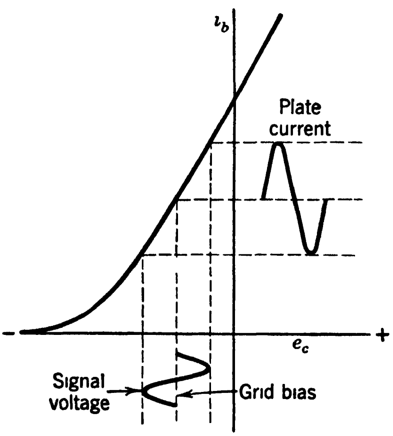


FIG. 4. Class A amplifier operation.

A *Class B* amplifier is one in which the grid bias is approximately equal to the cutoff value, so that plate current flows for approximately one-half of each cycle of input signal voltage. The operation of a Class B circuit is illustrated in Fig. 6 showing how only one loop of the signal wave appears in the plate-current circuit. The signal voltage applied to the grid of a Class B amplifier is usually of a much larger value than that applied to the grid of a Class A stage. The applied signal voltage may be so large that, during part of the positive half-cycle, the grid is operated at a positive potential with respect to the

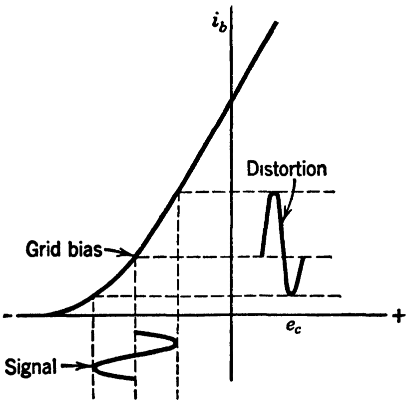


FIG. 5. Distortion in a Class A amplifier by excess bias.

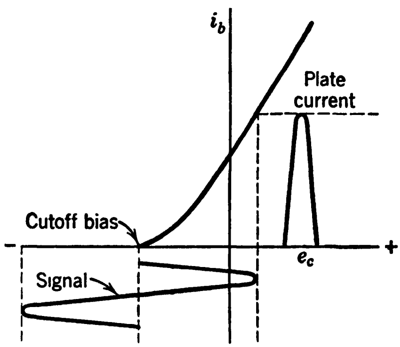


FIG. 6. Class B operation.

cathode. During this period some electrons will be attracted to the grid, constituting a grid current.

The large distortion present in the output of a single-tube Class B amplifier stage depicted in Fig. 6 may be reduced by using two tubes in a push-pull amplifier circuit (Fig. 7). The two tubes operate from and into transformers with a center-tapped winding so that each tube operates only on alternate half-waves and thus furnish an output wave form having small distortion. The tube *A* functions on the positive half-cycle of the signal voltage, as shown in Fig. 6, while tube *B* with strong negative signal on grid is inactive. For the next half-cycle (negative with respect to the first) the grid of tube *B* becomes less negative and amplifies while tube *A* is inactive. The amplified current loops pass through the primary of the output transformer in opposite directions and thus induce a complete wave form of voltage in the secondary of the output transformer. In addition to supplying a complete voltage wave in the output, the push-pull circuit greatly

reduces the second and all even harmonics which may be introduced within the tube (due to nonlinear characteristic) so that for a given power output much less total distortion is introduced.

Class B amplifiers have an efficiency of 50 to 60 per cent which means a reduced value of plate dissipation and an increased power output for a given power input. They are used generally where it is desired to develop a relatively large power output in the load circuit. Single-tube, Class B amplifiers are never used for audio-frequency amplification. However, they can be used successfully in radio-frequency

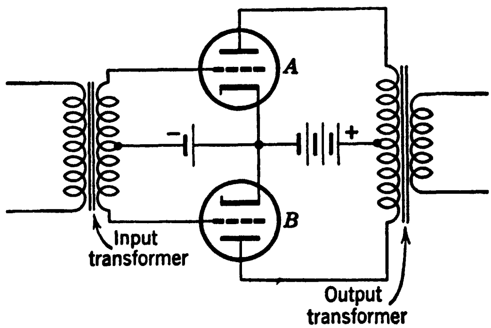


FIG. 7. Push-pull amplifier circuit.

amplifier stages having a parallel-tuned *LC* circuit for the plate load. The parallel-tuned circuit (antiresonant) previously discussed has the ability to store power. Thus during the positive half-cycle of the single tube Class B, the inductance *L* conducts current and the capacitor *C* becomes charged. Then during the negative half-cycle the capacitor discharges through *L* with current flowing in the reverse direction. Thus the load current flows through *L* in both directions and a complete cycle of amplified voltage will be induced in a second coil coupled to *L* (transformer action). This so-called flywheel effect of the tank circuit occurs only when its resonant frequency matches the frequency of the signal voltage.

Class AB amplification occurs when the amplifier tube is biased to operate part way between Class A and Class B. This form of amplification is used whenever it is desired to compromise between the high fidelity of Class A and the relatively high efficiency of Class B. Push-pull amplifier circuits are generally used for Class AB amplification to balance out the second harmonic current.

A *Class C amplifier* is one in which the grid bias is appreciably larger than the cutoff value so that plate current flows for appreciably less

than one-half cycle of the applied a-c signal voltage. The operation of a Class C amplifier is illustrated in Fig. 8. Here the grid-bias voltage is twice the cutoff value, which represents common practice in many type C amplifiers. The plate current flows only during the positive peaks of the applied signal voltage; hence the signal must be much larger than the cutoff bias to give satisfactory values of plate current. The parallel-tuned or tank circuit is used as a plate load to supply the negative loop of load current as suggested under type B amplification. The advantage of Class C amplification is its high efficiency which may reach 75 to 80 per cent. Class C operation is never used in audio-frequency amplifiers because of its high degree of distortion but it has wide application in radio-frequency circuits.

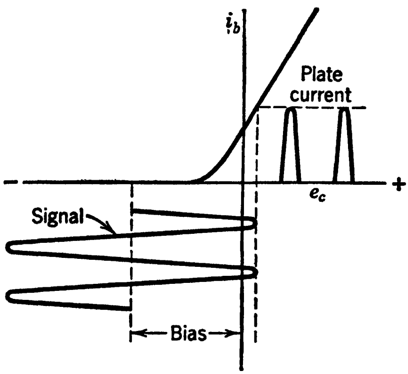


FIG. 8. Class C operation.

Interstage Couplings. The output of each stage of an amplifier must be coupled (connected) to the input of the succeeding stage. Several types of circuit are used for performing this function, with the choice determined by the tubes used, the frequency involved, the power transformed, and other factors. The simplest coupling circuit is illustrated in Fig. 9 and is known as direct coupling. In this circuit the plate of

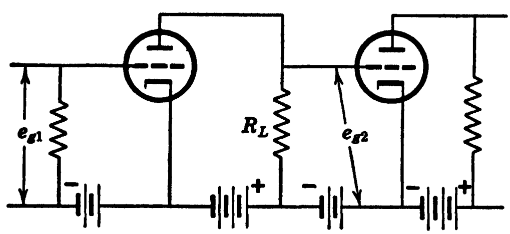


FIG. 9. Direct coupling of an amplifier.

stage 1 is connected directly to the grid of stage 2 and the cathode of stage 2 is connected to the positive side of the B supply for anode of stage 1. This means that the cathode (and tube parts) of each succeeding stage will be maintained at a higher potential above ground and that a separate plate-supply voltage must be provided for each

stage. This is the type of coupling used for d-c amplifiers which find applications in control equipment. The d-c amplifier is difficult to operate with a high fidelity because any change in cathode emission, grid-bias voltage, or plate voltage of one stage will be amplified in succeeding stages and will produce an effect similar to a change in the original signal voltage. This amplifier is also costly to construct because of the separate sources of power supply required for each stage.

Three common forms of interstage coupling are shown in the circuits of Fig. 10. The first stage of this circuit is known as resistance

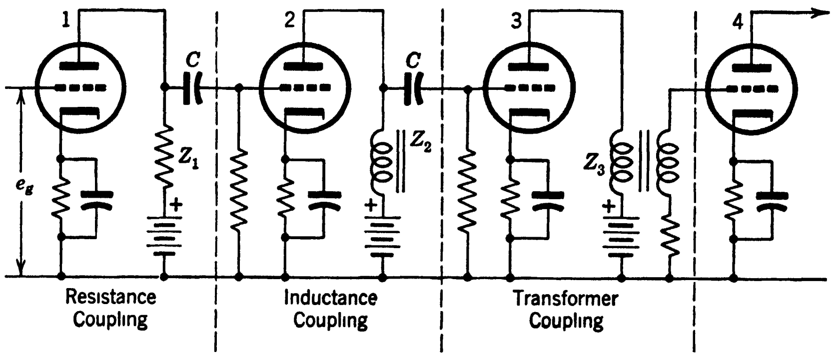


FIG. 10. Types of coupling.

coupling (or RC) since the amplified signal is developed across a resistor in the plate circuit. This signal voltage is applied to the grid of the second stage through the blocking capacitor (or coupling capacitor) C. The resistor in the grid circuit of the second stage provides a d-c path for the bias applied to the grid. Cathode-grid bias is used on all tubes in this circuit. The blocking capacitor plays an important part in the interstage coupling because it applies the amplified signal voltage across the resistor Z_1 to the grid of the next stage and simultaneously blocks the flow of direct current from the plate circuit of the first stage to the grid of the second. If this capacitor should break down or develop leakage, some of or all the d-c plate voltage of stage 1 would appear on the grid of the second and cause distortion in the output. The coupling between the second and third stage of Fig. 10 uses an inductance having a high impedance Z_2 at the signal frequency instead of a load resistance. This high impedance is used to match tubes such as pentodes which have a high a-c plate resistance. The general theory of operation is the same as for the first stage.

The third interstage coupling of Fig. 10 uses a transformer to connect

the output of the third tube to the input of the fourth. The primary of this transformer is the plate load for the third tube, while the secondary develops the signal voltage to be applied to the grid of the fourth tube. The use of the transformer permits additional voltage amplification within the transformer itself through the use of more turns on the secondary than on the primary. Impedance matching between the output and the input of the coupling is facilitated through the use of the transformer.

The three types of interstage coupling illustrated in Fig. 10 are used for audio-frequency amplification. Similar circuits are used for radio-frequency amplification except that tuned circuits are incorporated into the output and sometimes the input stages. These tuned or tank circuits are used since only a narrow band of frequencies are transmitted in radio-frequency circuits. Examples of the tuned types of coupling are shown in Fig. 11.

The amount of gain produced in an amplifier may be controlled in several ways. The simplest method to control the gain (or volume) is shown in Fig. 12 wherein a potentiometer form of variable resistance determines the magnitude of signal applied to the grid. A second method of gain control utilizes a variable-mu tube having a high degree of curvature near the cutoff of the i_b-e_c curve. A variation of the grid bias determines the point on the curve where the grid signal is applied and thus controls the gain of the tube. In radio-receiver circuits this method of gain control is made automatic through the use of a circuit wherein the magnitude of the amplified signal in one of the final stages controls the grid bias of a tube in some preceding stage, thus giving what is known as automatic volume control.

In the process of amplification the input signal may become distorted because of nonlinear characteristics of the tube and the circuits involved. If the relative amplitude of parts of the signal wave are changed, the result is called *amplitude distortion.* If the amplification varies with the frequency of the signal, the name *frequency distortion* is applied. A shift of the phase relationship between different frequencies during amplification is called *phase distortion.*

The characteristics of an amplifier can be changed by taking a portion of the output power and feeding it back to the input along with the input signal. Assuming that the output wave form is a replica of the input, it should be possible to combine it with the input signal satisfactorily, provided that the phase relations are correct. Obviously, two methods of combination can be made. In one method the transformed signal will be in the proper phase to strengthen the regular

input signal. This method is known as *regeneration* and was used in the early forms of single-tube radio receivers. Instability and critical operation of these circuits caused their use to be discontinued. However, this principle of regeneration is basic in the operation of oscillators which are to be treated in the succeeding article. The second

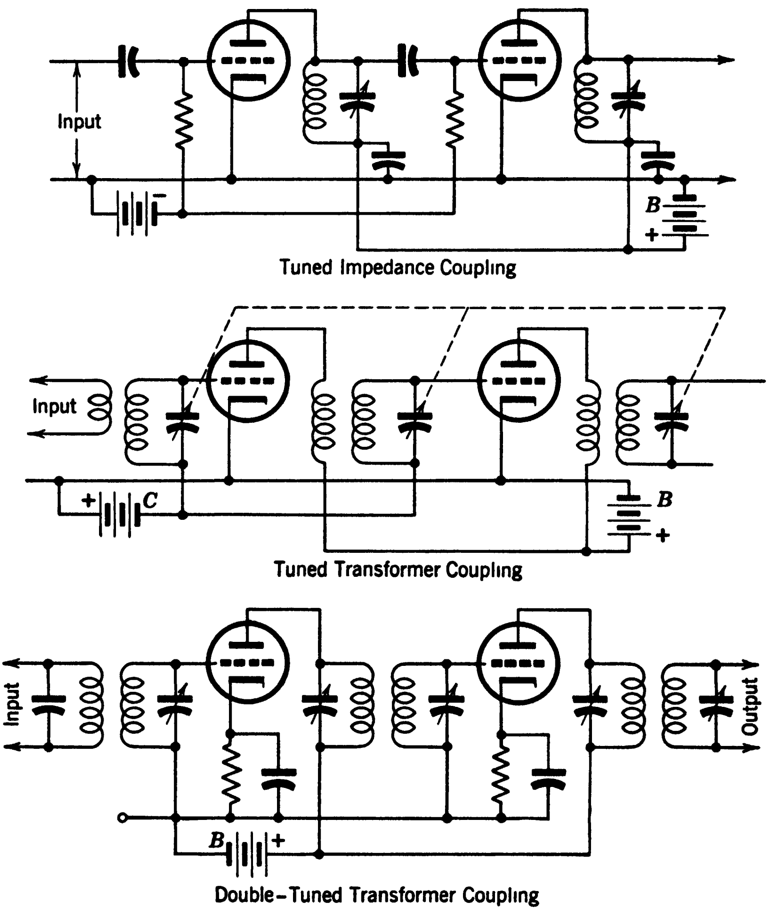


FIG. 11. Tuning for radio-frequency amplifier circuits.

method of combination for feedback is to apply the return voltage 180 degrees out of phase, or in opposition to the input signal. This application is known as degeneration or *negative feedback* for amplifiers. The application may be made within a single stage but is generally carried from the output of an advanced stage to the input of an earlier stage. Two advantages may accrue from the use of negative feedback. One

advantage is the automatic control of the gain in the amplifier which follows because any tendency toward an increase in gain in the intervening stages results in a greater feedback which restrains such gain. A second advantage of feedback is that distortion which arises within the amplifier stages is fed back in a negative sense and hence tends to cancel or reduce such distortional components.

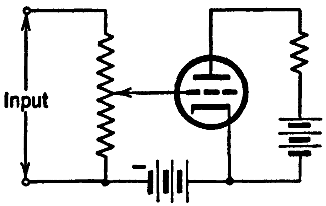


FIG. 12. Simple gain-control circuit for an amplifier.

Amplifiers have a very wide field of application. They make up a large part of the circuits in radio transmitters and receivers (both audio and radio frequency), they form a vital part of long-distance telephone circuits, and they play an important part in industrial electronic control in circuits to be given subsequent treatment.

2. Vacuum-Tube Oscillators. A vacuum-tube oscillator is a combination of a tube and a circuit which serves to convert direct current into alternating current. Such an oscillator functions by utilizing the amplifying power of tubes and the resonating property of the tuned *LC* circuit. A simple oscillating circuit is shown in Fig. 13. If switch *S* is thrown to the left the battery will charge the capacitor by removing electrons from the top plate and storing them on the lower plate. This

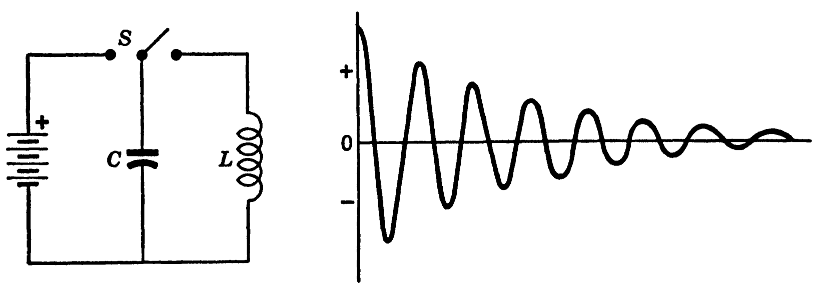


FIG. 13. Elementary oscillatory circuit.

action stores energy in the electric field of the capacitor. Now turn the switch *S* to the right and the capacitor will discharge through the inductance *L* with electrons moving from the lower plate back to the upper. The rising current through *L* will store energy in the magnetic field surrounding it. When *C* becomes discharged the energy of its charge will have been transferred to the magnetic field of *L*. This

stored energy in L will continue the flow of electrons and begin to charge C with a reversed polarity. This process continues until all the energy in the magnetic field has been transferred to C. At this point C begins to discharge again with a reversed direction of electron flow. Obviously, when C has released all its stored energy to the inductance L, the latter will have acquired energy to recharge C with the same polarity as originally provided by the battery. Now the circuit is restored to its original condition and is ready to repeat the process. If both the inductance and capacitance were without resistance or any form of loss, the resulting ideal circuit would continue to oscillate indefinitely. Such ideal circuits cannot be realized and some resistance is always present. Such resistance will reduce each swing of current as illustrated in the right view of Fig. 13. The larger the value of circuit resistance, the more rapidly the oscillations will be damped out. A freely swinging pendulum will have its oscillations damped out with time (like right view of Fig. 13). In a clock the pendulum is kept swinging with a uniform stroke by adding enough mechanical energy to each stroke to supply the losses due to friction and windage. In a similar manner, the LC circuit of Fig. 13 may be made to continue oscillations of uniform magnitude by adding the necessary electrical impulse at each swing. The LC circuit has a natural period or frequency which is determined by its resonant frequency f_r, where

$$f_r = \frac{1}{2\pi\sqrt{LC}}$$

and the magnitude of this frequency can be controlled through changes of L and C.

Any amplifier circuit will reproduce and amplify a signal voltage of the frequency applied to it. Now if a part of the output is fed back to the grid in the proper phase relation, the amplifier will be self-excited and under proper conditions will oscillate. A simple circuit to accomplish this result is shown in Fig. 14. In this circuit, coil L_p in the plate output is coupled inductively to coil L of the tuned LC circuit which, in turn, establishes the signal voltage applied to the grid. When the switch S in the plate circuit is closed, current flows in the plate-cathode circuit through L_p. The rising flux in L_p threads coil L inducing a voltage which charges C. When the plate current reaches its normal value, the energy in the magnetic field of L overruns, charging C to a higher potential. After reaching a peak level, C discharges into L and the LC tuned circuit oscillates at a frequency determined by its

resonant frequency. The oscillating grid signal causes the plate current to oscillate which, in turn, feeds back enough energy to overcome the losses in the LC tuned circuit. Thus the entire circuit of Fig. 14 becomes a self-excited amplifier.

The conditions necessary for sustained oscillations in a self-excited vacuum-tube oscillator are:

1. The feedback of power from the plate circuit to the grid must have a phase reversal of 180 degrees.
2. The power fed back must be sufficient to supply the losses in the grid input.

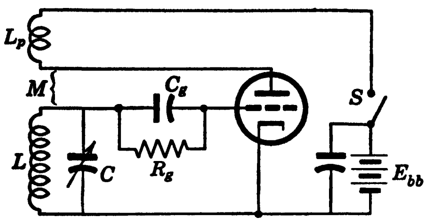


FIG. 14. Tickler coil oscillator circuit.

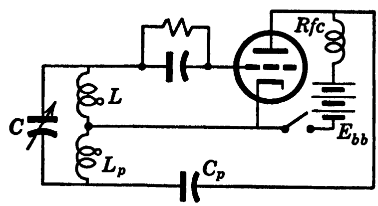


FIG. 15. Hartley oscillator circuit.

3. A tuned circuit LC is generally used in either the output or input or both to establish a resonant frequency. (Exceptions to this statement are given at the end of this article.)
4. Oscillators that are operated Class C (for high frequency) should have all or part of the d-c grid bias furnished by a grid leak. If a fixed bias is used, no initial plate current will flow and oscillations cannot start.

Obviously, many circuits can be designed to produce a vacuum-tube oscillator. Four of the more important ones will now be given.

One form of the *Hartley* oscillator circuit is given in Fig. 15. Here the coupling of the load to the input is combined with the tuned LC coil so that the a-c component of the plate current flows through L_p to the cathode. The radio frequency is kept out of the plate supply by the *Rfc* choke. The frequency of oscillation is controlled by the variable condenser C, while the signal voltage applied to the grid is adjusted by the cathode tap between L and L_p. The capacitor C_p prevents any short circuit of the plate supply through the tuning coil.

The *Colpitts* oscillator circuit of Fig. 16 is similar to the Hartley circuit since a pair of capacitors C_1 and C_2 replace the series inductances L and L_p, giving a capacitive type of feedback. Tuning is accomplished by varying the inductance of L though it is possible to use a

fixed L and to vary C_1 and C_2 by a ganged variable control. The grid-bias resistor must be connected directly to the cathode to provide the d-c bias.

An *electron-coupled* oscillator circuit is illustrated in Fig. 17. This circuit uses a tetrode connected so that its cathode, control grid, and screen grid act as an oscillator similar to the Hartley circuit of Fig. 15,

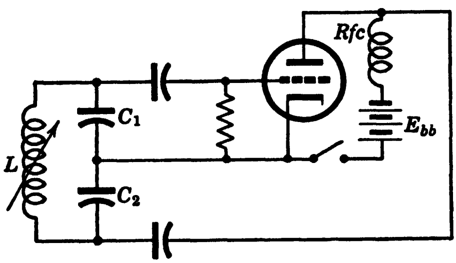


FIG. 16. Colpitts oscillator circuit.

while the plate circuit of the tetrode serves in the capacity of an amplifier, thus giving a greater output capacity. The coupling between the two circuits and the two functions is the electron stream within the tube, and hence the term electron-coupled oscillator. In this oscillator a tuned circuit is used in both the input and output circuits. Since the screen grid is held at radio-frequency ground potential and also serves

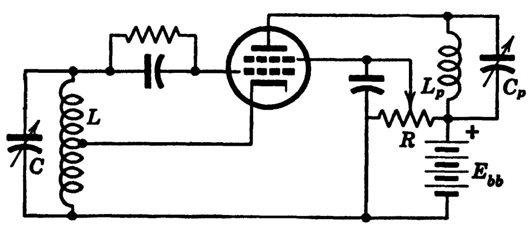


FIG. 17. Electron-coupled oscillator circuit.

as a shield between the two circuits, this oscillator is very stable and load variations have little effect on frequency change. Another factor that aids the stability of the electron-coupled oscillator is that an increase in screen voltage will decrease the frequency, whereas an increase of plate voltage will increase the frequency. Thus a proper adjustment of the tap on the resistor R will make the frequency independent of supply-voltage variations.

The frequency of oscillations generated by the tickler-coil, Hartley,

and Colpitts oscillator circuits is affected considerably by changes in load, supply voltages, and temperature. While the variation in frequency is small in electron-coupled oscillators, it is sufficient to be objectionable in broadcast transmitters, telephone carrier systems, and similar applications. Where precision frequency control of a constant frequency is necessary, crystal-controlled oscillator circuits are employed.

Certain crystalline substances such as quartz, Rochelle salts, and tourmaline exhibit mechanical and electrical properties known as the piezoelectric effect. Thus, if a mechanical force is applied to one of

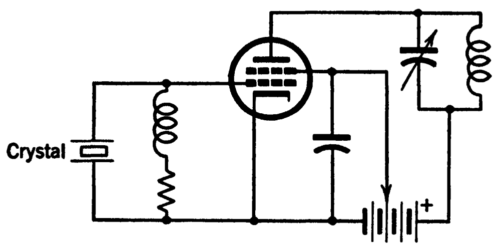


FIG. 18. Crystal oscillator circuit.

these substances, a voltage is developed. Conversely, if a thin slab of the substance is connected to a source of alternating voltage, it changes its physical shape and produces mechanical vibrations. Thin pieces are cut from quartz crystals for use in crystal oscillator circuits. When such a crystal starts vibrating at its resonant frequency, it will take only a small force of the same frequency to obtain vibrations of a large amplitude. The mechanical resonant frequency of a crystal depends chiefly on its thickness. When an alternating voltage is applied to a crystal that has the same mechanical frequency as the applied voltage, it will vibrate, and only a small voltage need be applied to keep it vibrating. In turn, the crystal will generate a relatively large voltage at its resonant frequency. If this crystal is placed between the grid and cathode of a vacuum tube and a small amount of energy is taken from the plate circuit and applied to the crystal to keep it vibrating, the circuit will act as an oscillator. The natural frequency of a crystal is critical (and precise). Thus if the constants of the oscillator circuits are properly adjusted the crystal will assure a precise frequency output.

A crystal-controlled oscillator stage using a tetrode tube is shown in Fig. 18. It will be noted that this circuit is similar to the electron-

coupled oscillator except that the quartz crystal and grid leak have replaced the tuned input circuit. The feedback takes place through the plate-to-grid capacitance within the tube.

The power output demanded of vacuum-tube oscillators depends on the application. For the majority of applications the oscillator is used merely as a source of a high-frequency signal which is fed into an amplifier to secure the required power. For such uses a power output from the oscillator of 1 to 5 watts is ample. This low power output is desirable because larger loads tend to affect the stability and frequency of the Hartley, Colpitts, and similar circuits. The electron-coupled circuit, however, does provide moderate power output with good stability. There are a few applications, such as high-frequency heating (to be covered later), where stability of frequency is unimportant but where a large power output is needed. For such applications circuits of the Colpitts, Hartley, or other types may be used with water-cooled tubes and circuit components of high voltage and current capacity so that a power output of from 1 to many kilowatts may be attained. In those circuits the load inductor or capacitor for the heating process may constitute all or a part of the tank circuit for the plate of the oscillator tube. For other applications the oscillator is loaded by an inductive or capacitive coupling to the plate circuit of the oscillator tube.

Oscillators employed in industrial electronics generally employ some combination of parallel *LC* for establishing the operating frequency. Some oscillators employed in other fields do not use such *LC* combinations. In one type of oscillator the 180-degree feedback from the plate circuit is secured by networks of resistors and capacitors *RC*. Each net (R and C) gives some angle of phase shift such as 60 degrees and a triple net will furnish the desired 180 degrees of shift. Another oscillator circuit employs the negative resistance characteristic to be found between two electrodes (such as screen and plate in tetrode) to produce oscillations. A third form of oscillator uses a vibrating reed (mechanical) to establish the frequency of oscillation.

3. Saturable-Core Reactors. The reactance of an iron-core inductance can be controlled by varying the saturation of its iron core. One method of applying this principle is shown in the construction and circuits of Fig. 19. The series line current passes from L_1 to L_2 through two series coils placed on the outer legs of a three-legged iron core. These coils are connected so that the flux which they produce circulates through the outer legs (magnetomotive forces opposed in central leg) and they are wound with turns of wire having a low ohmic resistance.

A second coil having a large number of turns and a high resistance is placed on the central leg of the magnetic circuit. If the second coil is not excited and a-c current passes through the series L_1L_2 circuit, a high reactance is presented by the resulting a-c flux in the iron core. Hence the a-c current is reduced by the inductive reactance to a very low value. The reactance of the device may be varied by passing a controlled d-c current through the second coil. Such current produces a d-c flux which moves through all legs of the iron core and at a suitably

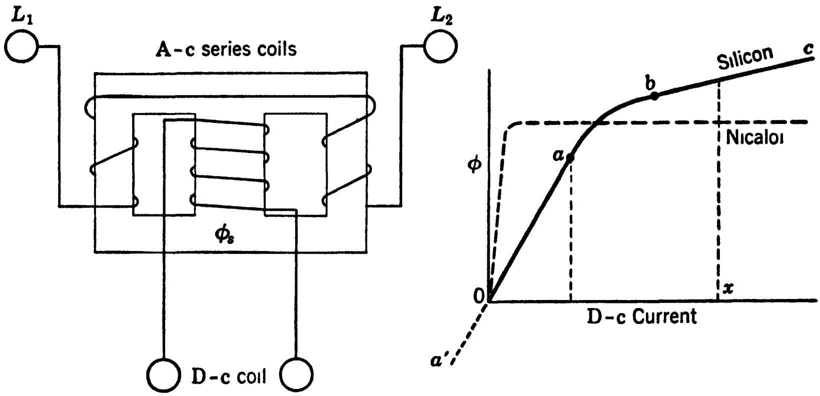


FIG. 19. Saturable-core reactor.

high value saturates the iron. The effect of saturation of the iron may be studied through the use of the magnetization curves at the right of Fig. 19. For zero applied d-c current the series coils produce flux which operates on the linear section $a0a'$ of the magnetization curve and high reactance results. Now if the d-c coil has a current of value ox, the iron core is saturated and any a-c current flowing will not produce any flux change (other than exists in an air core) and the reactance will be nearly zero. Over the range of a to b on the silicon saturation curve, a varying amount of flux change and resulting reactance will exist in the series line circuit. In this region of nonlinearity, harmonics will be introduced into the a-c circuit.

Some saturable reactors are constructed with the four-legged core shown in Fig. 20. These are more expensive to build but give a more efficient control because of the improved coupling between the d-c and a-c magnetic paths. Where it is desirable to reduce the d-c control power to a minimum the magnetic core may be constructed of Nicaloi (see curve, Fig. 19), which will reduce d-c excitation and give a very sharp change at the saturation bend.

Saturable reactors are designed so that flux changes arising from the current in the load circuit do not encircle the d-c control winding and thereby induce high voltages in it. The necessary magnetic balance to achieve this result is obtained in the three-legged core by connecting the load windings in series so that their magnetomotive forces are opposed in the central leg. For the four-legged core the d-c winding surrounds both the load windings so that the resultant magnetomotive force (sum of equal up and down values) is zero. The load windings may be connected in series or in parallel. The series connection gives a shorter time response (measured in cycles) and a higher voltage drop. It is generally used for controlling loads of small magnitude. The parallel connection is generally employed for controlling loads, using relatively large currents (½ to 75 kilovolt-amperes) and where a slow time response of 1 to 3 seconds is not objectionable.

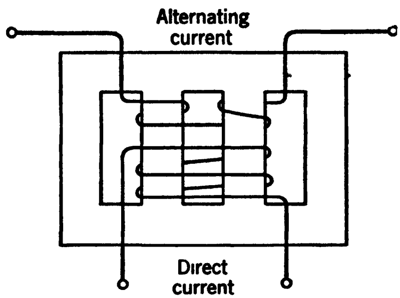


FIG. 20. Four-leg core for saturable reactor.

Saturable reactors are used commercially in a wide range of sizes for controlling a few volt-amperes up to many kilovolt-amperes. The d-c

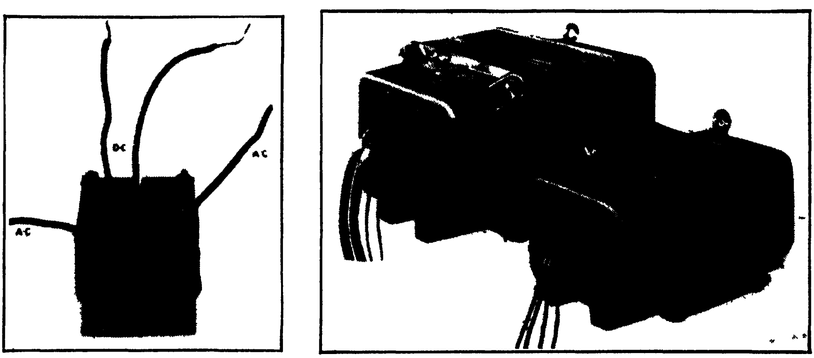


FIG. 21. Saturable reactors: *left*, small size for thyratron phase-shift circuit; *right*, 10-kilovolt-ampere and 4-kilovolt-ampere units. (Courtesy General Electric Company.)

control power varies from a few microwatts up to 20 or 30 watts. It is possible to control a small saturable reactor with a thermocouple or a photovoltaic cell, although the problems of shielding from stray magnetic fields becomes difficult in such applications. The d-c control

current is usually a few milliamperes output from a vacuum tube or a thyratron. Since the saturable reactor gives a large amplification of control power, it is sometimes termed a *magnetic amplifier*. Three commercial saturable reactors are shown in Fig. 21.

Two disadvantages to the use of saturable reactors are (1) slow response due to the inductance in the d-c circuit and (2) the production of harmonics arising from the nonlinear shape of the magnetization curve.

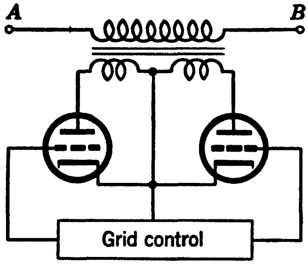


FIG. 22. Series impedance transformer control.

4. Series Impedance Transformers. The series impedance transformer varies the reactance in a series circuit like the saturated-core reactor but operates on the principle of reducing the flux in an iron core instead of forcing it to saturation. In any typical transformer on open circuit the reactance of the primary is very high. If the secondary coil is short circuited the resulting current produces a magnetomotive force which opposes the flux in the core and reduces the primary reactance to near zero. For intermediate loads and currents in the secondary the reactance of the primary will vary between maximum and minimum values. This principle of control may be utilized electronically by the circuit of Fig. 22. Here the current per-

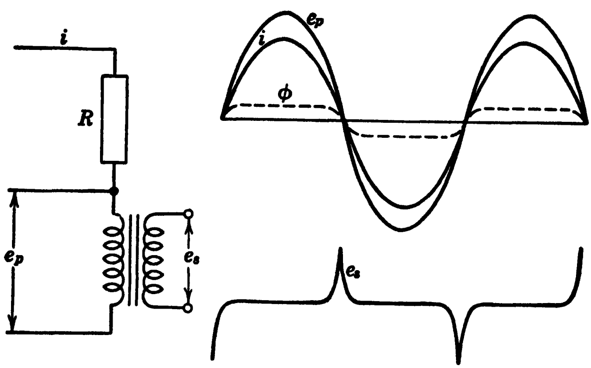


FIG. 23. Voltage and flux waves in a peaking transformer.

mitted in the secondary of a transformer placed in *series* with an a-c line is controlled by a triode (vacuum or gaseous) full-wave rectifier. A suitable grid control applied to the triodes either manually or automatically will vary the series impedance throughout the desired range.

5. Peaking Transformers. A peaking transformer is a low-capacity transformer operated with an over-saturated iron core for the purpose of producing peaked voltage waves in its secondary. The primary of this transformer is connected in series with a resistor (or inductance) across an a-c supply as indicated in Fig. 23 (left). The resulting wave forms of voltages, current, and flux are depicted on the right of the figure. With a limited amount of iron in the transformer core, saturation comes early in the current wave, giving a long flat-topped flux

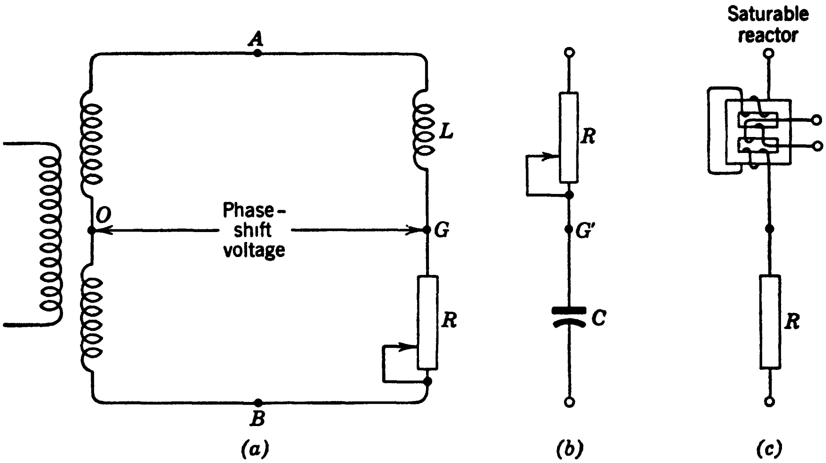


FIG. 24. A-c bridge circuits for producing phase shift.

wave. The flux changes in the iron core occur in a short interval of time in the current reversal zone, thus giving the peaked secondary voltage waves e_s. Peaking transformers are used for firing trigger tubes such as thyratrons and ignitrons.

6. Phase-Shift Circuits. The phase shift of voltages on single-phase circuits is readily controlled by the use of an a-c bridge (Wheatstone type). Such a circuit (Fig. 24) uses a transformer with a mid-tapped secondary to form two bridge arms AO and OB having equal induced voltages. The other two arms are made up of an inductance and a resistance as in a or a capacitor and a resistance as in b. A voltage phase shift from coil AB exists across the bridge points O and G. The variation in phase shift is produced by a variation of R, L, or C. The explanation of the shift is shown in the vector diagrams of Fig. 25. The voltage drop across AB must be equal to the vector sum of the drops across the arms L and R (or R and C) in series. In each case the IR drop and IX drop will be at right angles to each other. A reduction

of resistance R will reduce the IR drop and through some increase of the current will increase IX. A change of L or C will have a similar result. Thus a variation of the magnitude of IR and IX arms of the impedance triangle will cause the point G or G' to swing on the arc of a circle AGB or $AG'B$. Since the midpoint of the transformer O is fixed, the voltage across OG is constant in magnitude but varies in

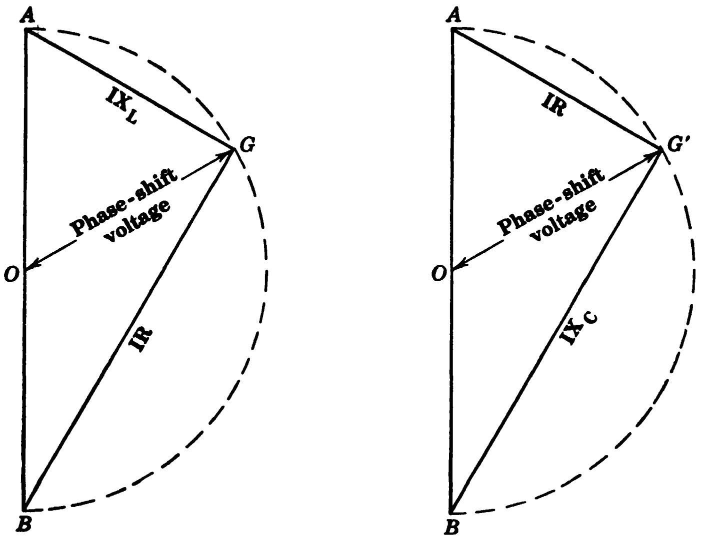


FIG. 25. Vector diagrams illustrating phase shift.

phase with respect to AB as G or G' swings around the arc. Theoretically, the maximum phase shift is 180 degrees, but the full value cannot be realized in a practical circuit.

The phase shift in the a-c bridge may be controlled manually through a variable resistor R or through a variable inductance (variometer). For some industrial applications the inductance L is varied by sliding an iron core within a solenoid. For electronic control the variation in the inductance is generally secured by replacing L with a saturable reactor (part c of Fig. 24). It is possible to use a series impedance transformer in place of the saturable reactor. A variation in phase shift can be secured also by varying the capacitance C, though this method has little practical use.

Phase shift for the firing of thyratrons can be secured effectively by a grid-bias control or by a combination of grid bias and an a-c voltage, as was pointed out in Chapter X.

Voltage phase shift for three-phase and multiphase circuits can be secured by the same principle as described for single-phase circuits though other methods are to be preferred. The simplest and cheapest method is to place a potentiometer (resistance) across one phase of a three-phase circuit as illustrated in Fig. 26. The phase of the voltage between C and X can be varied through a range of 60 degrees by moving point X along the resistor R. Phase shift can be secured by manual control of a phase shifter (similar to Fig. 27). This device consists of an a-c motor stator having a three-phase winding and a rotor wound with a similar three-phase winding. The phase of the voltage induced in the rotor relative to the stator will be determined by the fixed position of the rotor.

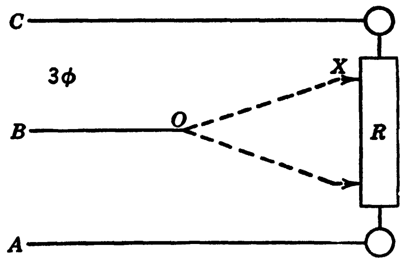


FIG. 26. Phase shift from a three-phase circuit.

7. Free-Wheeling Circuits. An inductance or choke coil connected in series with the output of a half-wave rectifier effectively limits the magnitude of the current and permits the flow of mere pulses of current. This limitation is often undesirable where magnetic relays are operated by thyratrons or vacuum-amplifier tubes. One circuit for overcoming this limitation is shown in part a of Fig. 28 where an inverted phanotron is connected across the inductance coil. The action of the circuit is pictured in part c of the figure. Without the phanotron the electron current passed by the thyratron will be limited to small pulses as shown by curves marked i because when the thyratron cuts off at the end of the positive half-cycle the current in the inductance must come to zero after some carry-over angle owing to the inductive effect. With the addition of the phanotron the energy stored in the magnetic field of L causes the electron current to continue flowing in the same direction through L by virtue of the circulating path provided by the

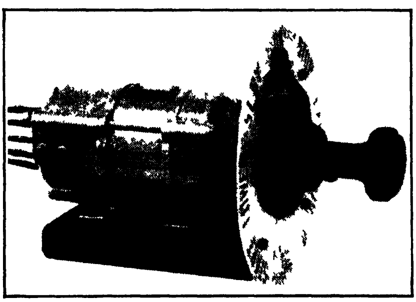


FIG. 27. Hand-operated selsyn transmitter and phase shifter. (Courtesy General Electric Company.)

phanotron and L. On the next positive wave of impressed voltage e the current rises to a higher value as shown by the full-line curve i_f. Within a few cycles the rectified half-wave pulses rise to a relatively high average magnitude. Since the current "coasts" through the phanotron on the negative half-waves, the circuit has been named the "free-wheeling circuit."

A useful circuit which performs a similar function is shown in part b of Fig. 28. This circuit may be used for relays operated by vacuum-amplifier tubes or small thyratrons. In this circuit the voltage drop across the relay charges the condenser on the positive half-waves, and

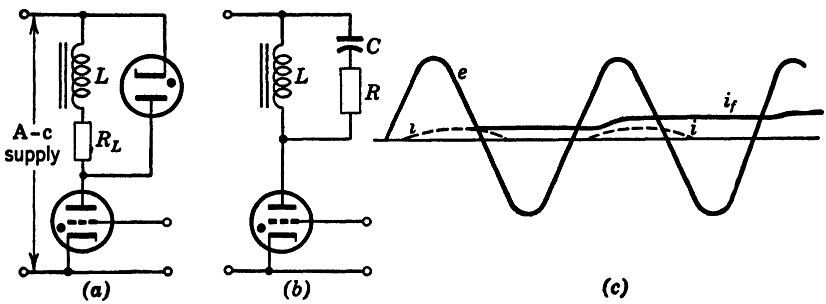


FIG. 28. Free-wheeling circuit.

the condenser discharges electrons through the relay (in same direction) during the negative half cycle. This action serves to hold the relay up and prevent it from chattering on the "off" half-cycles. The resistor R protects the cathode of a gaseous tube from peak current surges into the condenser but may be omitted with vacuum tubes.

8. Rectifiers. The rectifier is a device for converting a-c power to unidirectional or d-c power and is the most common and important of all components of industrial electronics and control. The unilateral characteristics of vacuum and gaseous tubes (diodes and triodes) which make rectifying action possible were covered in the treatment of electron tubes. Another important device for rectification is the blocking-layer type of rectifier exemplified by the copper oxide, selenium, and magnesium-copper sulphide rectifiers.

9. Filters. An electric filter is a circuit having frequency discrimination. In industrial applications the filters are used for (1) smoothing out the current and voltage waves found in the output of rectifiers, (2) producing pass bands for a-c frequencies, or (3) elimination of a band of frequencies. The smoothing type of filter is explained on page 285, Chapter XII. Discrimination against all except a narrow

range or band of frequencies f_r is provided by a simple LC circuit in series with the line as shown in part a of Fig. 29. The same LC circuit placed across or in parallel with a line will short out and prevent from passing the narrow band of frequencies corresponding to its resonant frequency. A narrow band of frequencies may also be eliminated by

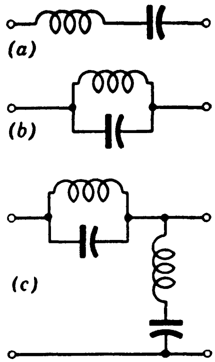


FIG. 29. Resonant filter sections.

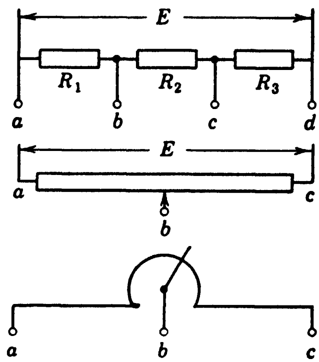


FIG. 30. Voltage-divider circuits.

the use of the parallel LC circuit of part b, Fig. 29, when placed in series with the line. A more effective filter network for the elimination of a narrow band of frequencies consists of the combination of parallel and series LC circuits as shown in part c of Fig. 29. In the field of

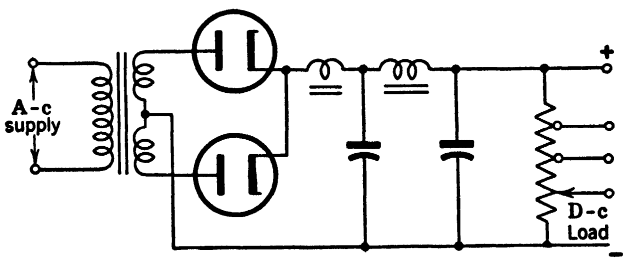


FIG. 31. Power pack.

communication, hundreds of types of filter circuits have been designed for various kinds of frequency discrimination.

10. Voltage Dividers. A voltage divider is a circuit for dividing a given voltage into two or more parts. Such division is readily attained by impressing a potential across a number of resistances in series as shown at the top of Fig. 30. Since the same current must flow through

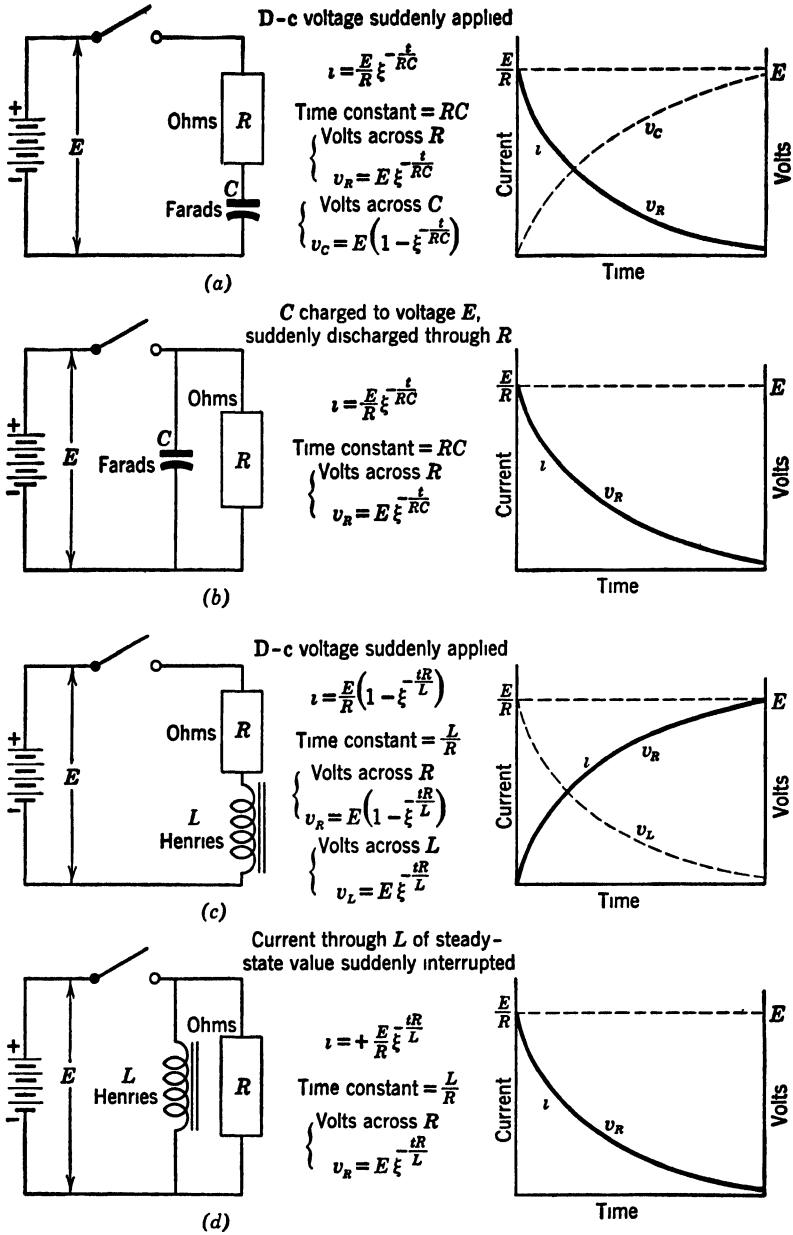


FIG. 32 Assemblies of timing circuits

the series, the drop in volts across each resistance *IR* is proportional to the resistance in ohms. If, in the top part of Fig. 30, *E* is 100 volts, R_1 is 4 ohms, R_2 is 6 ohms, and R_3 is 10 ohms, the respective voltages appearing across the resistors will be 20, 30, and 50 volts. A wide range of voltage can be obtained by using a moving contact *b* on a long linear resistor or on a resistor in a circular form as depicted in center or lower circuit of Fig. 30. The latter forms of voltage dividers are called potentiometers.

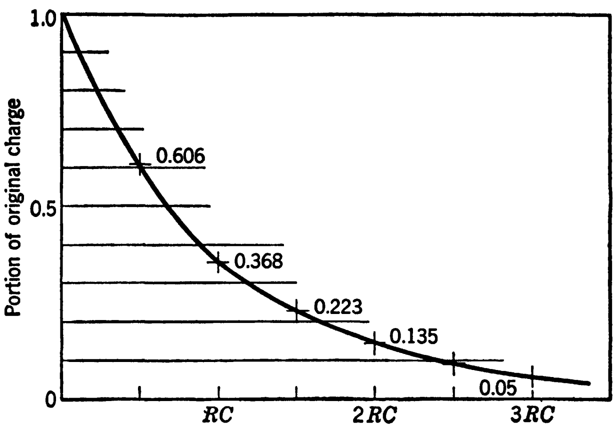


FIG. 33

A *power pack* is a combination of the three preceding components—a rectifier, a smoothing filter, and a voltage divider. This combined circuit, shown in Fig. 31, has a very wide application as a source of d-c supply in all kinds of communication and industrial electronic equipment.

11. Time-Control Devices. The timing of operations in industrial control may be performed by electric circuits, magnetic devices, and mechanical devices. The time required for transient current and voltage changes in inductances and capacitors is very useful in timing operations. Four common combinations of *RC* and *RL* are shown in Fig. 32 together with the equations and curves of transient currents and voltages. It is recommended that the student review these relations carefully. Combinations of *R* and *C* are the most useful circuit groupings. If, for example, a d-c source is used to charge a capacitor in parallel with a resistor (Fig. 32*b*) and the switch to the supply is opened, the capacitor will discharge at a definite rate. The rate of discharge is determined by the time constant *RC* of the circuit (see Fig.

33). After an elapsed time equal to the time constant RC, the voltage will fall to 37 per cent of its initial value. Similarly, after a second interval equal to RC, the voltage across the condenser will decay to 37 per cent of the value held at the end of the first period. Since this rate of voltage decay is exact and reliable, this form of timing circuit is very useful. A simple formula for this circuit is time (seconds) = resistance (megohms) $\times$ capacity (microfarads), for decay to 37 per cent initial value. Thus, in Fig. 33, if R is 1 megohm, C is 1 microfarad, and E is 100 volts, the voltage will fall from 100 to 37 in 1 second. Doubling the value of either R or C will double the time re-

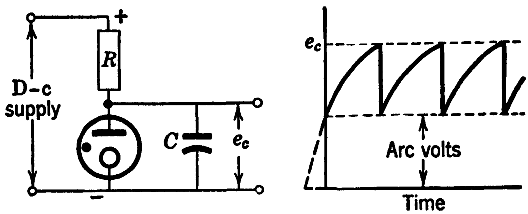


FIG. 34. Simple relaxation oscillator circuit.

quired for the same voltage change, thus giving simple timing control.

The shape of the voltage decay curve with time can be controlled by the substitution of some nonlinear device in the place of a resistance. For example, a pentode tube properly used in place of R will give a linear decay curve.

The rise of voltage across a condenser while it is charging may be used to give another form of timing circuit. Thus, in Fig. 34, a resistor and capacitor in series are connected across a source of d-c supply and a glow tube is placed in parallel with the condenser. When the d-c voltage is first applied, current flows through the resistor R into the condenser and begins to charge it. The voltage across the condenser ($e_c = C \times$ charge) rises until it reaches the "firing" or discharge value of the glow tube. When the glow tube fires it takes whatever current is passed by R plus a heavy discharge current from the capacitor. The voltage e_c across the capacitor falls very rapidly until it reaches the extinction point—the minimum ionizing potential across the glow tube. Now the glow tube ceases to conduct and the capacitor starts to charge again. When the capacitor voltage reaches the tube firing potential, the process is repeated. Thus this simple circuit, known as a relaxation oscillator, will produce a series of sawtooth voltage timing waves (Fig. 34, right) which are useful in oscilloscopes

and similar devices. A thyratron can be substituted for the glow tube in the relaxation oscillator circuit to give more power and a wider range of control.

The rise and decay of transient currents and voltages in series and parallel circuits containing R and L are similar to those in the RC circuits, as shown in parts c and d of Fig. 32. The time constant for the RL circuit is L/R, and the time required for a rise or drop to the 37 per cent value can be increased by increasing L or decreasing R. One useful application of the RL circuit is made in the construction of

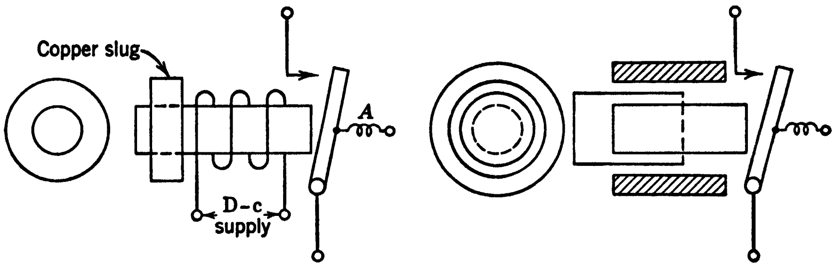


FIG. 35. Types of slow-acting relays.

slow-acting relays. In Fig. 35, left, a relay is shown which consists of a d-c coil and a heavy copper ring or "slug" placed on an iron core. When the d-c coil is excited the core becomes magnetized and attracts the armature A and closes a circuit. When the d-c circuit to the coil is opened the flux in the core begins to decay but the decrease of flux within the copper ring gives a change in the flux linkages, causing an emf and resulting current which opposes the change in flux. The L/R time constant of the copper ring determines the rate of decay of flux. Since only a small amount of flux is necessary to hold up the armature when in contact with the iron core, there will be some delay before the armature is released. The length of time delay can be controlled in a fixed design by the axial length and position of the copper ring. The time delay can be made variable by the use of movable copper sleeve (tube) placed between the iron core and coil as shown in Fig. 35 (right).

Many mechanical devices are used for timing operations. The clock (both spring type and electric) has uses for "on" and "off" operation. Small electric motors with contacts operated through reduction gearing offer good control for timing intervals from 30 seconds to several minutes. Electrically heated bimetallic strips will give a slow motion suitable for timing operations. The oil-filled dashpot is a hydraulic

device which will give a controllable time delay as well as serve as a damper for antihunting.

12. Constant-Voltage Circuits. Control circuits sometimes require a constant d-c voltage independent of supply-line potential. Constant voltages may be attained by connecting a voltage regulator or glow tube in series with a high resistance, as shown in the circuit of Fig. 36. The glow tube has the property of maintaining a constant voltage drop between cathode and anode for a wide range of current through the tube. The voltage drop across the glow tube depends on its design and rating. Three or more constant potentials may be secured by placing two or more tubes in series as in Fig. 36.

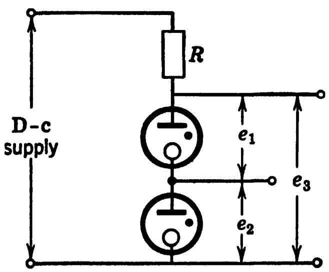


FIG. 36. Circuit for providing constant voltages.

13. Nonlinear Resistors. Some substances have a nonlinear resistance under changes of temperature, current, and voltage. The ballast tube or current-regulator tube of Fig. 37 contains a helical coil of wire having a very high positive temperature coefficient of resistance. The rise in resistance with temperature is great enough to hold the current through it nearly constant over a fairly wide range of impressed voltage, as suggested in Fig. 38. Ballast lamps are used to protect cathode heaters and other circuit parts from transient currents and the variations of voltage supply.

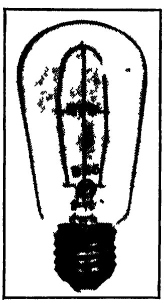


FIG. 37. Ballast or current regulator tube (Courtesy General Electric Company.)

Thyrite is the trade name for a substance that offers a very high resistance to a low impressed voltage but decreases its resistance with rise in voltage (Fig. 38). For example, doubling the voltage will increase the current sixteen times. This property makes Thyrite an ideal material for protecting tubes and circuit parts from insulation breakdown arising from high inductive transient voltages (Thyrite shorts out the circuit).

The copper oxide rectifier has a low resistance to current flow in one direction but a very high resistance to current in the opposite direction. This unilateral characteristic is useful for rectification of alternating current but also has applications in industrial control circuits for blocking current flow in one direction (Fig. 38).

Thermistors—the word is a contraction of "thermal resistors"—are members of a family of solid variable conductors that have a negative coefficient of resistance with rise in temperature. Thermistors are made of various combinations of manganese, nickel, cobalt, and other metallic oxides. In some types the resistances may be doubled with a temperature decrease of as little as 30 degrees F.

The metallic oxides that form the heart of Thermistors are milled, mixed thoroughly, and then made into desired shapes by forming beads on fine wires, by pressing into disks, or by extruding as rods. Subsequent processing and firing give the elements a hard, ceramic structure to which connecting leads are attached. Bead Thermistors are usually mounted in small sealed bulbs which are evacuated or gas filled; disk Thermistors are soldered to mounting plates and provided with a protecting finish; rod Thermistors are coated with glass.

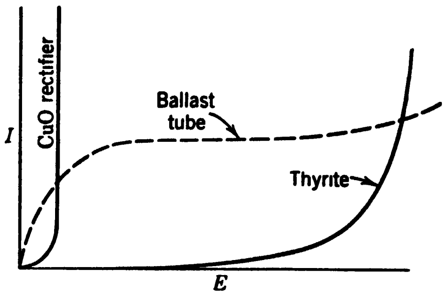


FIG. 38. Nonlinear resistance characteristics.

Thermistors may be used wherever temperature variations exist or can be produced. The variations in temperature may be brought about in three ways: (1) externally, by changes in surrounding air, water, etc.; (2) internally, where the current through the Thermistor changes its temperature; (3) indirectly, by means of a heating coil surrounding the Thermistor element.

Thermistors have numerous applications in telephone and broadcasting service. Their desirable properties should find applications in the industrial control field. The temperature charasteristics of these Thermistors are shown in Fig. 39 and one commercial form of these units is illustrated in Fig. 40.

14. Electronic Contactors. Electron tubes are often used as electronic contactors. Any of the vacuum tubes, such as the triode, tetrode, or pentode, will block a current flow in the plate-cathode circuit when the grid is biased to cutoff. A positive shift of the grid potential will permit a flow of current for "off" and "on" switching operations. In a similar manner, the so-called trigger tubes, the thyratron, ignitron, and grid-glow tube, can be used as an electronic switch via grid control. The thyratron is adapted to control a moderate

amount of power in an "on" and "off" switching operation. When it is desired to pass complete cycles of alternating current, two thyratrons may be employed as in the circuit of Fig. 41. Each thyratron serves to pass alternate half-cycles subject to individual grid control. The ignitron is used to control a large amount of power involving up to hundreds of amperes. One circuit for switching a heavy a-c load is

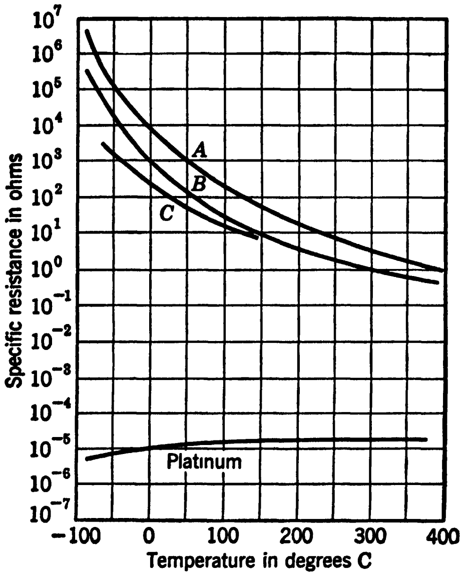


FIG. 39. Temperature-resistance characteristics of thermistors compared to platinum.

shown in Fig. 42. Two ignitrons, A and B, are employed so that one or the other is available for conduction in each direction. The igniters are fired by separate gaseous diodes, a and b, for the "on" and "off" control through the switch S. Operation may be traced as follows. With S open, line voltage cannot be impressed across either diode a or b and the ignitrons cannot be fired. Close switch S and assume that the line voltage (sine curve) is rising in positive direction on the anodes for tubes A and a. Then the cathodes of A and a will be negative so that, when the voltage equals the ionization potential of tube a, it conducts electrons from the lower side of the load through the mercury pool and the igniter of ignitron A to the cathode of a, and thence back to the line. This "shot" of current passing through the igniter of A fires ignitron A which then conducts a full wave of current having

a magnitude determined by the load. When the voltage reverses on the succeeding half-cycle the anodes of tubes B and b become positive and the cathodes negative so that b conducts and, in turn, fires B for the reversed wave of current. Thus the circuit of Fig. 42 will act as a closed switch to conduct a-c current as long as S is closed. When S is opened the main line will be opened the first time the a-c impressed voltage passes through zero. The substitution of some automatic control circuit in the place of switch S will give automatic control.

The gaseous diodes (phanotrons) of Fig. 42 may be replaced by blocking-layer rectifiers such as copper oxide or selenium. In some applications it is desirable to control the magnitude of the current as well as to provide "on" and "off" switch operation. This action can be secured by firing the ignitrons via thyratrons instead of phanotrons or blocking-layer rectifiers. A circuit for performing this function and another using copper oxide rectifiers is shown in Fig. 14, page 339. A commercial electronic switch employing ignitrons is shown in Fig. 15, page 340.

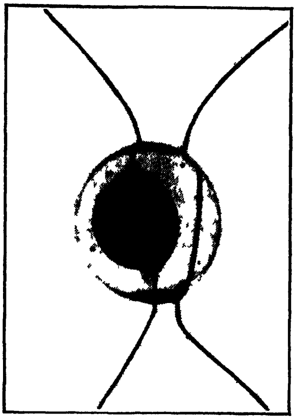


FIG. 40. Thermistor. (Courtesy Bell Telephone Laboratories.)

15. Long-Tailed Pair. One form of balanced or stabilized amplifier circuit used in industrial control applications is shown in Fig. 43. Two like triodes or other multielectrode amplifier tubes $T1$ and $T2$ (often contained in a single enclosure) are connected to equal load resistors $R1$ and $R2$. The common point for their cathodes is fed through a single cathode resistor $R3$ which has a large magnitude (say 100,000 ohms), giving rise to the term a *long-tailed pair*. The grid of tube $T2$ has a potential which is fixed with respect to the d-c supply by the resistors $R4$ and $R5$. The input to the amplifier is applied between the grids at points a and b, and the output appears between the plates of the tubes. If the input voltage is zero, both grids are at the same potential and the tubes will conduct equal currents which will be limited by the rise of potential across the common cathode resistor $R3$ and the drop of potential across equal load resistors $R1$ and $R3$. For balanced conditions both tubes will conduct the same current, and zero potential difference will exist across the output at points c and d. Now if a voltage is introduced into the input which makes a positive with respect

to b, tube $T1$ will conduct more current and the rise in current through $R3$ will produce a rise of potential across $R3$, which tends to reduce the

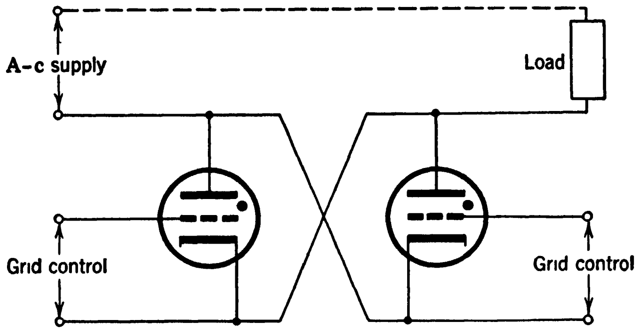


FIG. 41. Circuit contactor using thyratrons.

difference in potential between the cathode and grid of both tubes. The net result is a rise of current in tube $T1$ and a reduction of current in

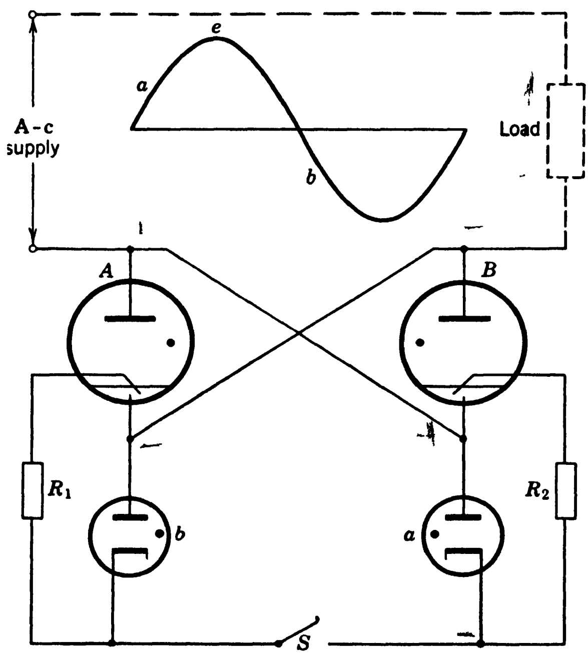


FIG. 42. Ignitron contactor using gaseous diode firing.

tube $T2$, which gives a difference in potential across the output cd. A lowering of the grid potential at a will produce an opposite effect. In

practice the sum of the currents through the tubes does not change much with variations of the input voltage.

An indicating instrument connected across the output of Fig. 43 converts the circuit into a vacuum-tube voltmeter. The substitution of the d-c windings of saturable reactors for resistors $R1$ and $R2$ makes possible some very useful control applications to be discussed in later chapters.

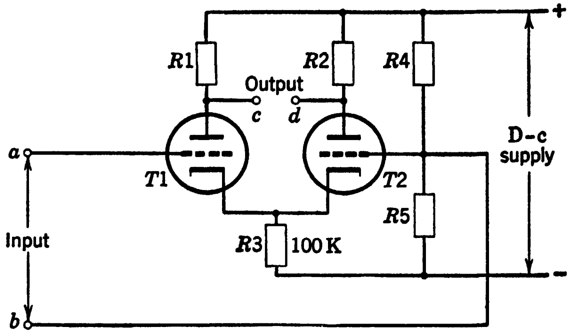


FIG. 43. Long-tailed pair circuit.

16. Light-Control Devices. The electronic components that function under changes in light are the phototube, the electron multiplier, and the photovoltaic cell. The theory of these devices was covered in Chapter VII and their applications will be covered in later sections.

17. Temperature-Control Devices. The measurement of temperature for applications in temperature control is accomplished by resistances, bimetallic strips, and thermocouples. Resistances constructed of wire having a high positive coefficient of temperature resistance may be used directly in circuits or may be used as one arm of a Wheatstone bridge to act as a pilot for control functions. Bimetallic thermal elements consist of two thin flat strips of different metals (usually brass and invar steel) welded together as a unit. The difference in the coefficient of expansion of the two bonded metals causes the strip to bend with changes of temperature. Bimetallic strips may be used in linear, spiral, or helical form (Fig. 44) and the resulting movement under changes of temperature may be used to trip mechanical devices or to open and close electrical circuits. A thermocouple is a junction of two dissimilar metallic conductors. When one junction of the two wires is held at a temperature differing from the other end of the wires, an emf is induced at the junction. This emf is directly proportional to the temperature difference and depends on the metals used for constructing the thermocouple.

18. Position-Control Devices. The function of position-control devices is to transmit motion by electrical means between two points that cannot be connected readily mechanically. This function can be accomplished through the use of either direct or alternating current. A simple d-c system for transmitting the motion of the float in the gasoline tank of an automobile to the indicator on the dash is shown

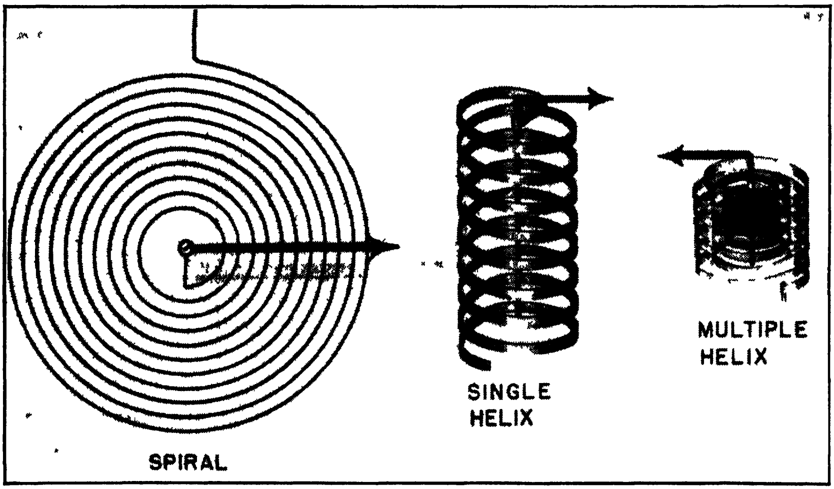


Fig. 44. Temperature response coils using bimetallic strip. (Courtesy Weston Electrical Instrument Corporation.)

schematically in Fig. 45. Two small solenoids, M and N, are placed with their axes at right angles and are connected in series across a 6-volt battery. A variable resistor R operated by the motion of the float is connected across coil M and the battery. A movable iron vane or armature A is pivoted at the intersection of the axes of M and N. When the tank is empty R is zero and N is short-circuited so that A lines up along the axis of M. As gasoline is admitted to the tank the motion of the float increases the resistance R, increases current through N, shifts the resultant field, and swings A to the right.

The d-c selsyn principle which is widely used on aircraft and some ships is illustrated in Fig. 46. The sender is a circular rheostat to which a d-c voltage is fed by a double slider. The receiver is a steel ring on which three coils are placed, each to span 120 degrees. For every position of the rheostat the current in the coils is so distributed that the flux cuts across the diameter of the steel ring at the same angle as the slider on the rheostat is placed. A magnetized vane pivoted

at the center of the steel ring lines up with the flux, thus following any motion given to the sender shaft.

The a-c selsyn system (also called synchro system) for transmitting motion electrically consists of two self-synchronous motors. The two motors are connected together by three wires as indicated by the circuit of Fig. 47*a*. These motors have a three-phase, wye-type of winding on their stator but have a two-pole, shuttle-wound rotor with its coil connected through collector rings to a single-phase, a-c source of excitation (Fig. 47*b*). In this system one machine is operated as a generator or transmitter and the other as a motor or receiver. When the rotor excitation circuit is closed, the a-c field of each rotor induces voltages in the three-phase winding on the surrounding stator. The three voltages induced in these phases are unequal in magnitude and are determined by the position (of rotation) of the rotor field. When the two rotors are in exactly corresponding positions, the voltages induced in the transmitter

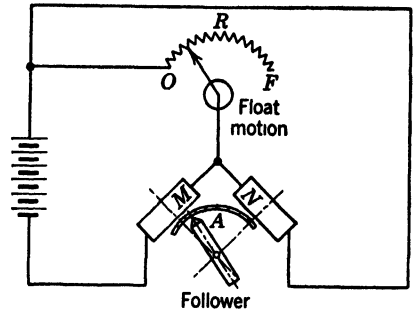


FIG. 45. Circuit for an automobile gasoline gauge.

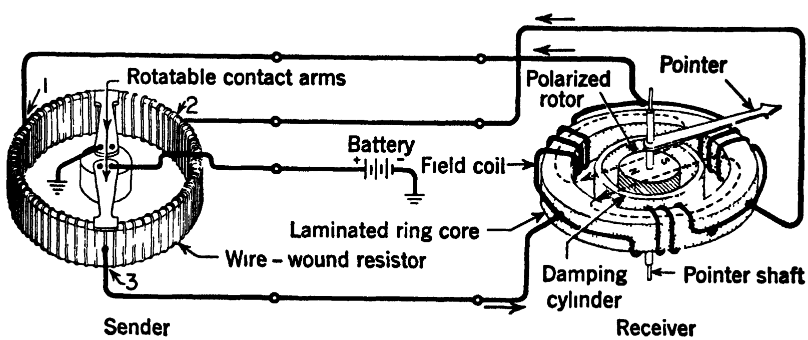
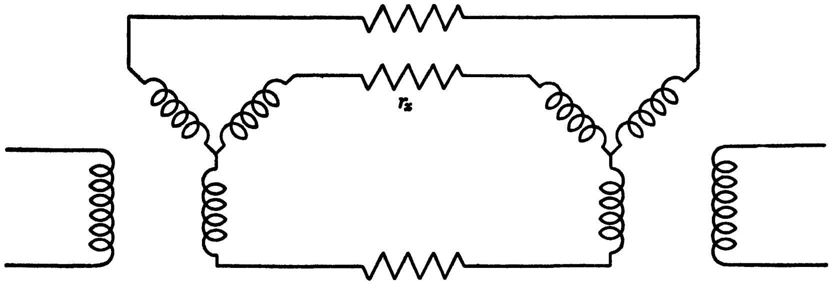


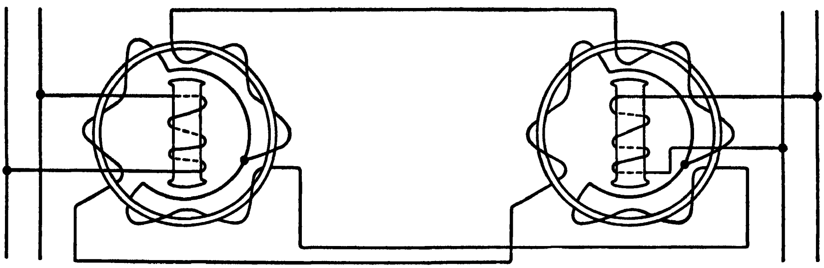
FIG. 46. Circuit for a d-c selsyn system.

stator are equal and opposite to those induced in the receiver stator; that is, they are balanced, so that no current flows in the winding of either stator. If, however, the transmitter rotor is moved from the original position, the induced voltages are unbalanced, current flows in the stator winding, and a torque is set up in both

rotors. Since the transmitter rotor is held in position, the receiver rotor moves under the developed torque until it occupies a position corresponding to the new position of the transmitter. A selsyn transmitter was shown in Fig. 27 and a high-accuracy selsyn unit is illustrated in Fig. 48. The latter figure displays the wound rotor for



(a)



(b)

FIG. 47. Circuit for an a-c selsyn system.

this selsyn unit which is equipped with an inertia damper to prevent hunting or oscillation.

Selsyn systems are used for indicating the positions of all kinds of remote mechanical equipment and also for the remote control of such equipment. They are made in a variety of sizes and with different degrees of accuracy to fit the application.

19. Rotary Amplifiers. New designs in the circuits and construction of small d-c generators have the property of high-gain power amplifiers analogous to some electronic amplifiers using vacuum tubes. The *amplidyne* is a widely used machine of this class. Its theory of action may be followed in the schematic circuit of Fig. 49. An input or control power is applied to a shunt field winding known as the control field. The excitation of this winding produces the field flux ϕ_{CF} which cuts the inductors of the revolving armature and induces an

emf across the brushes cc'. If the brushes cc' are short circuited, a relatively large current will flow in the armature inductors and produce an armature cross field ϕ_{SC} in the direction indicated. This will be a relatively strong cross field (armature reaction) which will be cut by the revolving inductors and thereby furnish an emf across the brushes bb'. The output or load for the generator is taken from brushes bb'. Any load current from bb' flowing through the armature inductors will produce a field ϕ_L that will oppose the control field (demagnetizing action) and tend to weaken the action of the control field. To neutralize this demagnetizing action, a load-compensating field (series) is placed on the same poles as the control field and carries the load current, producing a field ϕ_C in such a direction as to oppose ϕ_L. With this compensation correctly effected, the output of the generator should respond to changes of input in the control field and with a high degree of amplification.

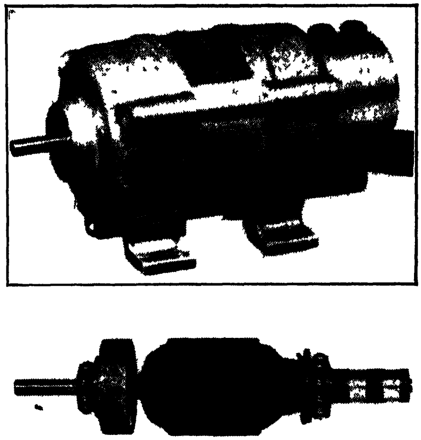


FIG. 48. Selsyn motor and wound rotor with damper. (Courtesy General Electric Company.)

The amplifying action of the amplidyne may be increased further by the addition of either the field S or S' shown in Fig. 50. These fields are placed on poles that produce flux along the same direction as the armature cross field ϕ_{CF}. Field S is a series field in the short-circuit path and S' is a shunt field connected across brushes bb'. While both types of field S and S' could be used on the same machine, there is no need for such combined use. The output of the amplidyne is controlled by one or more factors. Each factor involved requires a separate control field and perhaps a separate winding. One factor may be a constant and the corresponding field is known as a neutralizing, standard, or reference field. A minimum of two and a maximum of six control-field windings are used on amplidynes. A second control or neutralizing field is shown as N on Fig. 50.

Most amplidynes operate as two-pole generators though the actual machine has four segmental poles, as shown in Fig. 51. Segmental poles N and N' and similarly S and S' act as one pole. The control

fields are wound on the individual poles and then connected so that N and N' serve as one north pole structure. The compensating or load fields may be wound on the individual poles or as a single winding surrounding two segmental poles. Commutating poles and windings are provided to give good commutation at the load brushes.

The amplification factor of the amplidyne is defined as the ratio of the volt-amperes output to the volt-amperes input to the control field.

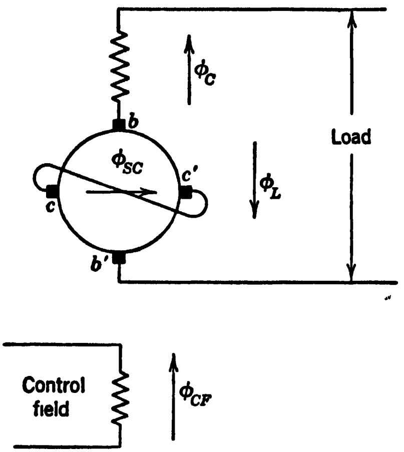


FIG. 49. Magnetic fields in an amplidyne.

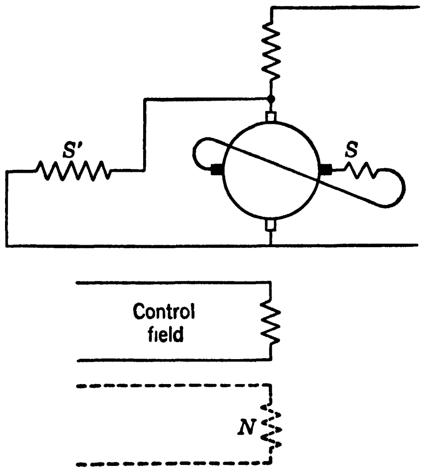


FIG. 50

The amplification is produced in two stages: first, from the control field to the quadrature field, and, second, from the quadrature field to the output. Amplifications varying from 100 to 100,000 are possible though the usual practicable range is of the order of 5,000 to 10,000. Amplification can be gained at the expense of time or rate of response. If amplification is carried too far it may be too slow to be useful, especially in the control of sudden changes in machines. The amplification factor increases with the size of the amplidyne because the required control watts do not increase as fast as the output watts. The important things to know in selecting or designing an amplidyne are the amplification factor and the rate of response. The rate of response decreases slightly as the size of the machine increases. The average rate of the rise of voltage in the output is of the order of 2000 volts per second. A commercial amplidyne is illustrated in Fig. 52.

Other machines bearing the classification of the rotary amplifier are

certain *multiple-field exciters.* Their theory of operation may be understood through a study of the several field-resistance lines of a typical d-c generator as shown in Fig. 53. For a field resistance of R_2 the no-load generated voltage will rise to R_2, the intersection of the saturation curve and the field-resistance line. For other values of field resistance other steady-state values of generated voltage may be found. There is some value for the field resistance for which the field-resistance line will coincide with and be tangent to the lower linear portion of the saturation curve. In Fig. 53 this is shown by the line OR. Since line OR does not intersect the saturation curve at any point where the lines coincide, the generated voltage could lie anywhere along this line of intersection and would not change unless disturbed by some change in field strength produced by some outside factor. But it will be very easy to shift the generated voltage up or down along this line by a change in field produced by a second control or signal winding of a few turns placed on the poles of the generator. Thus the control generator should have at least two separate fields, one designed to produce a field-resistance line coinciding with the linear part of the saturation curve, and the second a light field for receiving the control change or signal.

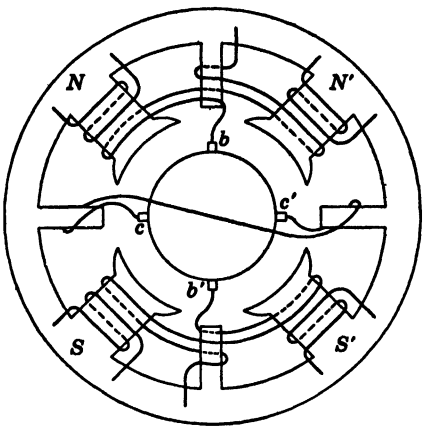


FIG. 51. Field circuits in an amplidyne.

In order to obtain best results in control, the following points should be observed in the design of the control generator. First, the main field should be designed so that with a suitable external resistor the resistance line may be made to coincide with the saturation curve. Second, the cross section of magnetic circuit should be made relatively large so that the saturation curve will be linear throughout the voltage range for which the generator must operate. Third, the iron used for the field should have a low hysteresis loss and the inductance of all parts of the circuit should be low so that a quick response to all current changes may be obtained.

The basic and essential adjustment of the control generator, the coinciding of saturation curve and field-resistance line, is brought

about by a self-excited field. This self-excitation may be secured by the use of a series field as shown in part *a* of Fig. 54, or a shunt field as in part *b* of the same figure. The first basic circuit is used in a commercial machine known as the Rototrol, while the latter circuit is used

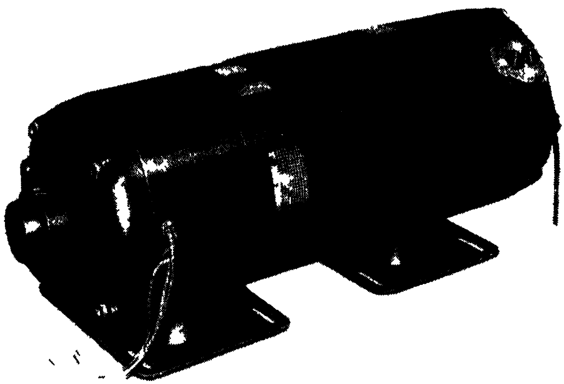


FIG. 52. One-kilowatt amplidyne. (Courtesy General Electric Company.)

in a similar machine under the name Regulex. In either circuit, the "resistor" is adjusted to give proper operation for the applications involved. In both circuits operation requires the introduction of the signal through the "control" field. For some applications a single control-field winding may suffice, but usually two or more field coils will be needed. When the multifield exciter is used for regulation of voltage, current, or speed, two fields are needed. One field, known as the standard, comparison, reference, or pattern field, is connected to some fixed and constant source of potential. The second field, known as the signal, pilot, or control field, is connected to the varying unit which is to be regulated. These two fields are connected in opposition either in series to one winding or to separate windings so that, when the variable has the desired value, the two fields neutralize each other and the voltage

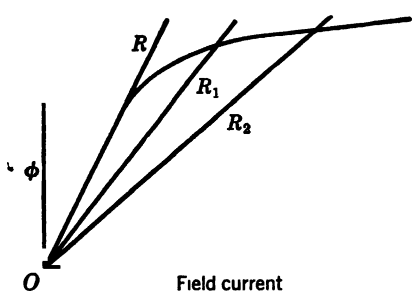


FIG. 53. Field-resistance lines of generators.

generated by the control generator is stationary. When the varying unit moves up or down it throws the reference and signal fields out of balance in such a direction as to cause a change in the generated voltage of the control generator which in turn instantly corrects the shift. In some applications all fields may be of the signal or control type, being connected to different varying factors and giving a composite change in the control generator output. A recent trend is to feed

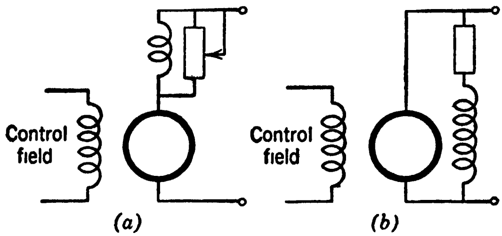


FIG. 54. Circuits for multiple-field exciters.

two electrical signals into an electronic amplifier so that their amplified difference operates a control field of the rotary amplifier.

Numerous applications of the multiple-field exciters are made on Ward Leonard systems for controlling speed, acceleration, deceleration, or torque of large motors. They may also be used where control is based on tension (paper, wire, etc.) or the "position" of some unit in an industrial application. These machines, like the amplidyne, can be used to control anything that is convertible into volts, amperes, or watts.

20. Antihunt Circuits. Antihunt circuits are designed to prevent oscillations in servomechanism systems and their theory of action becomes more intelligible in connection with servo theory. They are explained on page 267.

PROBLEMS

1. A triode used in a simple amplifier circuit has an amplification factor of 7.5 and a plate resistance of 6000 ohms. If the load is a resistance of 9000 ohms, what is the voltage gain of the amplifier? How can this gain be increased?

2. A saturable reactor near saturation carries a flux of 400,000 lines in its iron core. If a 115-volt, 60-cycle alternating current is supplied to the reactor and load through 200 turns on the series coils, what is the voltage across the reactor if the d-c voltage coil carries zero current and the resistance of the series coils is neglected?

3. Construct a magnetization curve for a silicon iron core as shown in

Fig. 19. Pass vertical axes (Y) through points 0, x, and the knee of the curve. Superimpose equal sine waves of current along these axes below the x axis and plot the resulting flux variations in the iron core. What will be the shape of the induced voltage wave across the inductance? In what region will harmonics appear in the current and voltage waves?

4. An RC timing circuit (Fig. 32*b*) consists of a 1-microfarad condenser in parallel with a resistance. What should be the value of the resistance used if the voltage across the condenser falls to 0.37 of its value during a period of 30 cycles of a 60-cycle circuit?

5. An RC timing circuit is to be designed to lower voltage across a condenser from 1.0 volt to 0.2 volt in a period of 5 cycles (60-cycle). If a 0.5-microfarad condenser is used, what will be the magnitude of the resistance? Repeat for a 15-cycle period.

6. A thyratron is to be fired by the voltage rise across the capacitor C in Fig. 32*a*. If the thyratron triggers when the voltage across the condenser rises to 0.63 of the battery value and the resistor has a value of 2 megohms, what must be the value of the capacitor to fire the thyratron in 3 seconds after the switch is closed?

7. In Fig. 32*c*, how long will it take the voltage across R to build up to 63 per cent of battery value if $L = 5$ henries and $R = 1$ ohm?

8. Assume in Fig. 34 that $R = 1$ megohm and $C = 1$ microfarad and that the glow tube breaks down when the voltage across it reaches 86.5 per cent of the direct current applied. What interval of time will transpire after the d-c potential is applied before the glow tube conducts?

9. What is the function of resistors R_1 and R_2 in Fig. 42? If the phanotrons are rated at 20 amperes maximum, what should be the value of R_1 and R_2 on a 220-volt a-c supply?

REFERENCES

GULLIKSON, F. H., and E. H. VEDDER, *Industrial Electronics*, John Wiley & Sons, 1935.

COCKRELL, W. D., *Industrial Electronic Control*, McGraw-Hill Book Co., Inc., 1944.

CHUTE, GEORGE M., *Electronics in Industry*, McGraw-Hill Book Co., Inc., 1946.

BENDZ, W. I., and C. A. SCARLOTT, *Electronics for Industry*, John Wiley & Sons, 1947.

Chapter XI

PRINCIPLES OF CONTROL AND SERVOMECHANISMS

The chief difference between man and other forms of animal life is that the animals rely upon muscular strength to satisfy the needs of existence, whereas man utilizes many forces of nature to accomplish his desires. Primitive man domesticated some animals to aid in his work. Later centuries saw the utilization by man of the forces of the wind and water in crude windmills, water wheels, and the sailing of ships. The nineteenth century witnessed the harnessing of the energy in steam and the beginning of the internal combustion engine. The first quarter of the twentieth century brought the development of electricity as the flexible source of energy for light, power, and communication. The second quarter of a century (1925–1950) is seeing a great advance in the automatic control of machines through the medium of electronics and other new electrical and mechanical devices.

Simple Control Systems. Simple control systems involve some manual actions which determine the operation of a machine. The machine being controlled may be driven by a mechanical prime mover, by hydraulic pressure, or by electricity. When electric energy is employed, the control may be an "on" and "off" action produced by a simple switch, as suggested in Fig. 1*a*. The simple switch may be replaced by some form of starting box or automatic push-button starter. Such devices function as starting or stopping mediums and are classed as simple controls.

The control of an electric motor may be continuous as indicated in Fig. 1*b*. Here the control is not of the "on" and "off" type, but the speed of the motor may be changed without stopping the motor by a variation of the series resistance R. A third form of simple control is suggested in Fig. 1*c*. Here sound produced before the transmitter at the left controls a much louder replica of this sound which issues from the loudspeaker at the right. In this control system the primary

sound (mechanical) varies the resistance in a microphone which, in turn, varies the magnitude of an electric current. The varying current fed into a transformer induces an alternating current in the secondary circuit which results in an amplified sound output from a loudspeaker.

In these simple control systems, the expenditure of a small amount of energy has served to control the delivery of a much greater amount of energy. This relationship is characteristic of all forms of control systems.

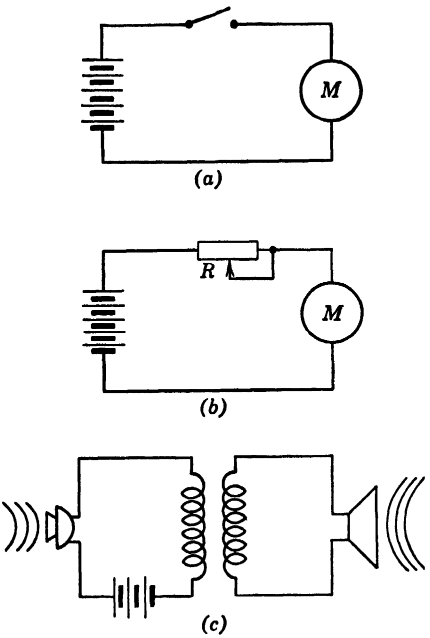


FIG. 1. Simple control systems.

Closed-Cycle Control System. A closed-cycle control system is one in which there is a mechanical or an electrical circuit interlock between the final controlled unit and the primary controlling unit. In electrical terminology this means a closed circuit and a feedback system. As an example, it is desired to maintain a small oven at a constant temperature for a crystal oscillator. The desired result may be accomplished by the circuit and arrangement shown in Fig. 2 where the oven is heated by an electric resistance heating element. A make-and-break contact operated by a bimetallic thermal unit is placed in the oven and connected in series with the heating element. The contact is adjusted on the temperature dial to give the desired oven temperature. When the temperature of the oven rises above the dial setting, the thermostat opens the electric circuit and permits the oven to cool slowly. After the inside temperature falls to the dial setting, the thermostat recloses. In this example the closed circuit or cycle is very simple and all active elements are located within the oven itself. The controlled factor is temperature and the primary controlling unit is the thermostatically operated contactor. In most closed-cycle or feed back systems several components are required in the complete system. These automatic closed-cycle systems are frequently termed servomechanisms. This terminology arises from the fact that, in the maintenance of the desired

factor or function, the controlled mechanism serves or is a *slave* to the primary controlling unit.

Servomechanisms. The following article is quoted from the September 1946 issue of *The Westinghouse Engineer* through the courtesy of the author, Dr. S. N. Herwald, and the publisher.

Servomechanism is a term appearing with increasing frequency in the engineering literature. Although the word is not new it confuses many, partly because it is not self-explanatory but more particularly because it applies not to any particular device but to a function. Servomechanisms can be con-

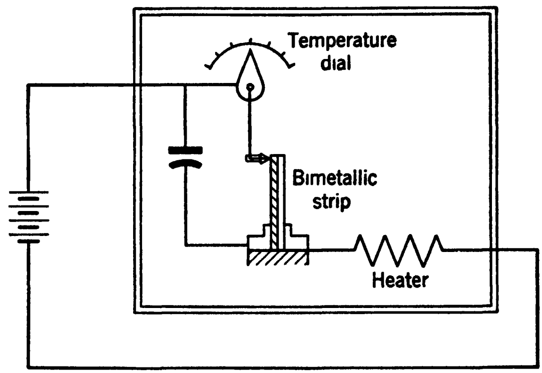


FIG. 2. Simple closed-cycle control system.

sidered a class of automatic regulators whose purpose is to keep a regulated quantity matched to a reference quantity. Usually some distance intervenes between the initiating device and the thing regulated. Moreover, in this matching process some amplification is involved.

The quantities most commonly regulated are the speed, position, or change in rate of speed of some machine. Simple mechanical devices to control liquid in a reservoir by the position of a valve in response to liquid level are servomechanisms. So also are temperature control systems using thermostats to control speed of motors or valves on a heating system or the motor of a refrigerator compressor. Hydraulic governors that control turbine speed to maintain a constant electrical frequency with load, devices to control motor speed on tandem strip mills and paper-making machines so that the moving strips of steel or paper do not become too loose or too tight between mill sections, tracer mechanisms that enable a cutting tool to reproduce the contours of a scale model, a stabilizing device to hold steady the gun of a moving tank, or the mechanism that moves a heavy gun turret around as the gunner sights on the target—these are but a few servomechanisms. Many operate electrically or electronically; others are mechanical, hydraulic, or even gas operated. Sometimes they are used in combination.

In keeping a regulated quantity matched to a reference quantity, four important characteristics must usually be incorporated within the servomechanism:

1. Fast response.
2. High accuracy.
3. Unattended control.
4. Remote operation.

These characteristics become more valuable as the difficulty of the control problem increases, although many simple problems, such as the control of

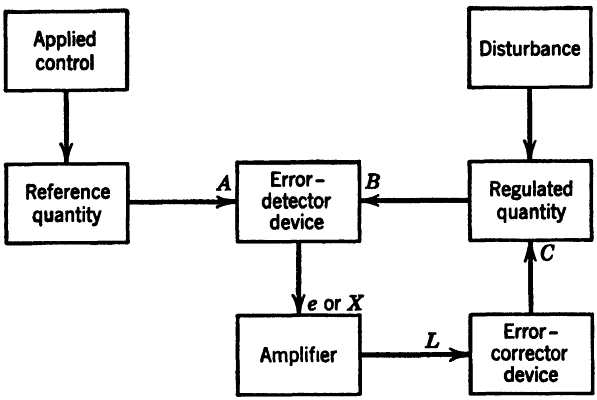


FIG. 3. The essential components of a servomechanism. (This figure reproduced with permission from the article "Forms and Principles of Servomechanisms," by Dr. S. M. Herwald, in the September 1946 issue of *The Westinghouse Engineer.*)

household heating by a thermostat and furnace, are everyday servomechanism applications.

A servomechanism consists of three basic elements: (1) an error-detecting device, (2) an amplifier, and (3) an error-correcting device. As indicated in Fig. 3, each serves a unique function in enabling the regulated quantity to stay matched to the reference quantity. The error-detecting device determines when the regulated quantity differs from the reference quantity. It then sends out an error signal to the amplifier, which in turn supplies power to the error-correcting device. With this power, the error-correcting device changes the regulated quantity so that it matches the reference quantity. The closed loop composed of the error detector, the amplifier, the error corector, and the regulated quantity is characteristic of all servomechanisms.

Any quantity can be regulated—voltage, speed, temperature, position, direction, torque, to mention a few of the most common. Any quantity also can be used as a reference quantity. The reference quantity need not be the same as the regulated quantity if the proper error-detecting device is employed. For

example, it is possible to make the speed of a turbine-driven blower vary with outside temperature, the speed of the blower being the regulated quantity and the outside temperature being the reference quantity. Most practical problems, however, concern the matching of two similar quantities such as regulated position against a standard position or regulated speed against a standard speed.

The error detector of Fig. 3 serves as a means of measuring both the reference quantity (A, Fig. 3) and the regulated quantity (B, Fig. 3) and of detecting when they are not similar, e or X being the measure of any difference. The measuring units within the error detector must be the same. For example, in the case of the blower governed by outside temperature, if the temperature is measured as a function of voltage by a thermocouple, then the speed of the blower is also measured as a function of voltage, perhaps by a tachometer. When the voltage representing the reference quantity, outside temperature, is compared to the voltage representing the regulated quantity, speed, the difference obtained is also a voltage. This small difference voltage is the error signal. It varies in magnitude with the amount the turbine speed is different from the required speed for a particular outside temperature, and it has a polarity depending upon whether the turbine speed is higher or lower than the required speed. A number of typical error-detecting devices are shown in Table 1, No. 1 to 10.

The amplifier of Fig. 3 is merely a means of using the signal of low power level (e or X, No. 1) provided by the error-detecting device to control the error-corrector device. Typical amplifiers are shown in Table 2. Familiar devices such as generators, valves, relays, and electronic tubes are included.

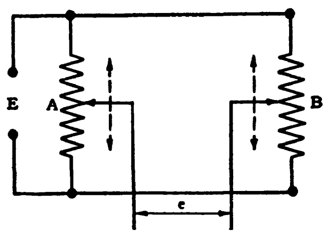
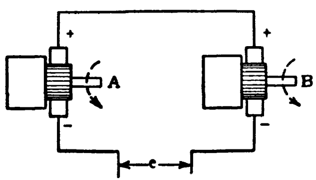
The error corrector of a servomechanism system actually does the work of regulating. The error-detecting device and the amplifier serve only as a means of controlling the error corrector so as to make the regulated quantity match the reference quantity. Several familiar types of error correctors such as electric and hydraulic motors, gas engines, and turbines are listed in Table 3.

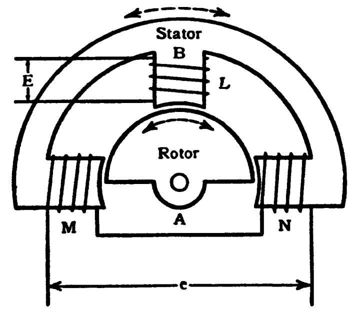
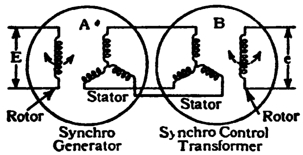
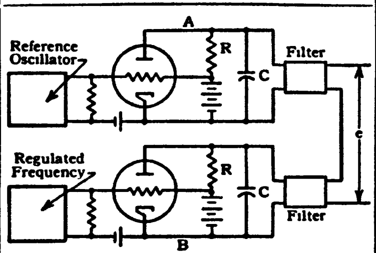
The disturbance shown in Fig. 3 represents load variations affecting the regulated quantity such as generator voltage drop with load for a voltage regulator, or temperature drop in a room with decrease in outside temperature for a temperature regulator.

Applied control as shown in Fig. 3 is controlled variation of the reference quantity in order to achieve desired changes in the regulated quantity. A good example of applied control to the reference quantity is the familiar one of variation of the reference temperature of a room thermostat. Both the applied control and the disturbance to the regulated quantity are independent. They are influences that upset the existing balance between the regulated quantity and the reference quantity. Servomechanisms readjust this balance continuously and automatically.

Error-Detecting Devices. The error-detecting device has many practical forms. Some of the representative types are given in Table 1. Each error de-

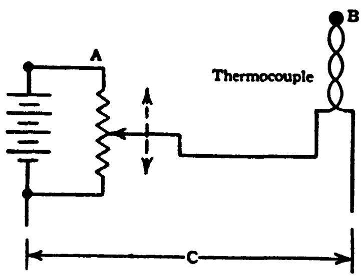
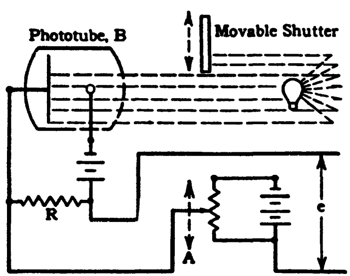
TABLE 1. ERROR-MEASURING DEVICES *

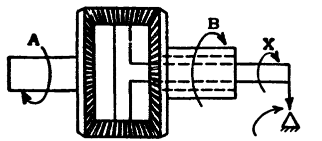
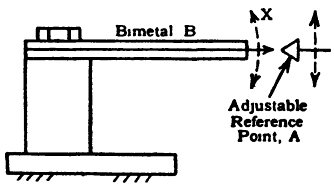
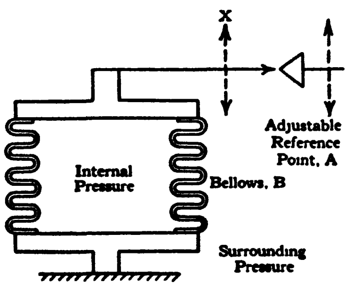
No.	Schematic Representation	Type	Main Application	Operation
1		D-c or a-c resistance bridge	Position control	Error voltage, e, appears when the position of the moving arms of the potentiometers A and B are not matched. The power source, E, is applied across both potentiometers. A measures reference position as voltage and B regulated position as voltage, their difference being e.
2		D-c tachometer bridge	Speed control	Error voltage, e, appears when speeds of tachometers A and B vary. A measures reference speed as a voltage and B regulated speed as a voltage. The difference between these voltages is e.

3		A-c magnetic bridge	Position control, particularly for gyro pickups where very small forces prevail	Error voltage, e, appears when relative positions of rotor A and stator B do not match. Rotor A measures reference position magnetically and stator B regulated position magnetically. Voltage E, across exciting coil, L, provides energy. When rotor covers unequal areas of each exposed stator pole (unbalanced magnetic bridge) pickup coils M and N have unequal voltages induced. Voltage difference is e.
4		A-c synchro-system (selsyn)	Position control where continuous rotation is desired	Error voltage, e, appears whenever the relative positions of the rotors of synchro-generator, A, and synchro-control transformer, B, are not matched. The reference position is measured by A as a magnetic flux pattern which is transmitted to the synchro-control transformer through the interconnected stator windings. If the rotor of B is not exactly 90 degrees from the transmitted flux pattern, e is produced.
5		Frequency bridge	Frequency control	Error voltage, e, appears when reference and regulated frequencies differ. Tube channel A produces a filtered saw tooth wave that gives a d-c voltage inversely proportional to the reference frequency. Tube channel B produces a similar voltage as a measure of the regulated frequency. The difference of these d-c voltages is e.

* This table is a condensation of material from the article on servomechanisms by Dr. S. N. Herwald, taken by permission from the September 1946 issue of *The Westinghouse Engineer*.

TABLE 1. ERROR-MEASURING DEVICES (*Continued*)

No.	Schematic Representation	Type	Main Application	Operation
6		Millivolt bridge	Temperature control	Error voltage, e, appears whenever the regulated temperature differs from the reference temperature. The regulated temperature is measured as a voltage by the thermo-electric effect of two dissimilar metals, B. The reference temperature is represented as a voltage from the battery-potentiometer source A. The difference in these voltages is e.
7		Phototube bridge	Position control by intercepting a light beam	Error voltage, e, appears when movable shutter is in other than desired position. Light reaching phototube, B, measures shutter position. This light is measured as a voltage by the phototube current variation. A reference position of the shutter is represented by the battery-potentiometer voltage. The difference of these voltages is e.

8		Mechanical differential	Position control and speed control	Displacement X appears whenever the relative reference and regulated positions change. Reference position is measured as an angle by one side of the differential, A, and regulated position as an angle by the other side of the differential, B. The difference in the two positions rotates the middle member of the differential giving displacement X.
9		Bimetal	Temperature control	Displacement X appears whenever the surrounding temperature and the reference temperature are different. The reference temperature is represented by the position of the adjustable reference point, A. The surrounding temperature is measured by the position of the bimetal strip, B. The difference in these positions produces displacement X.
10		Bellows	Pressure control and temperature control	Displacement X appears when surrounding and reference pressures differ. Reference pressure is represented as position by adjustable point, A. Surrounding pressure is measured by the bellows as a position. Difference in these positions produces displacement X.

tector, regardless of whether it is electrical, hydraulic, or mechanical, performs three distinct operations:

1. Measurement of the reference quantity. This element is shown by A in the schematic representation of Table 1.

2. Measurement of the regulated quantity. This element is shown by B in the schematic representation of Table 1.

3. Indication of amount of error. This element is shown by e if the output is electrical and by X if the output is mechanical in the schematic representation of Table 1.

Analysis of the d-c tachometer bridge, No. 2, Table 1, serves as a typical example. Tachometer A measures the speed of a reference, as a voltage. The speed of the device being regulated is measured by tachometer B as a voltage. The voltage outputs of tachometers A and B are opposed. A voltage difference, e, appears only when the speed of B differs from speed A. Thus a definite error indication is established.

The features of an error detector that most interest the designer are: energy required to measure reference quantity, accuracy, size and reliability.

The greatest amount of design ingenuity can be exercised in the selection of an error-detecting device. Special arrangements to decrease operating forces or to increase accuracy are used. Typical of these is the use of a dual synchro-system to increase accuracy.* In this arrangement one synchro-system, such as No. 4, Table 1, is geared at unity to the reference and regulated angular positions. A duplicate synchro-system is geared to the reference and regulated angular positions at some higher gear ratio, which is usually an odd number, such as 15 or 31 to one. When the angular discrepancy is large, the unity system takes control and through the amplifier and error corrector brings the angular positions into approximate agreement. At this point the vernier or high gear-ratio system assumes control of the amplifier and error-correcting device and very accurately brings the regulated position into correspondence with the reference position. The high-ratio or vernier system cannot be used alone for it would have as many correspondence points as its gear ratio above unity. The peculiar traits of mechanical, electrical, or hydraulic error detectors can also be used to advantage. As an example, magnetic saturation might be used to get a particular variation of error voltage with error. Similarly, a non-linear mechanical linkage might be used.

Amplifiers. Some typical amplifiers are shown in Table 2. The broad similarity of all the amplifiers of Table 2 is apparent when one considers that each contains a "gate" element, G, which controls power flow from the power source, P, to the load, L. Each of these gates is a low-power or force element when compared to the power they control. For example, the contact of No. 1, Table 2 requiring ounce-inches to operate may control many watts to the load. For high amplifications single amplifiers can be cascaded. A series composed of an electronic tube, No. 4, Table 2, a relay, No. 2, Table 2, and a motor-

* Selsyn system, see page 241.

operated throttle, No. 8, Table 2, would provide tremendous amplification. Microwatts into the tube could control thousands of horsepower. In servomechanism applications, the load, L, taken from the final amplifier always represents the error-corrector device. The difference in the power available as an error signal and power required to correct the regulated quantity determines the amount of amplification necessary.

Many devices such as the contact, throttle, or valve have their power input determined by the time required for operation. The clutch, for example, requires inch-pounds in a certain time to control horsepower. Therefore in the column for approximate power-amplification factor of Table 2, a time factor, t, is included. Thus the faster one operates the clutch, the larger the power source required and the less its power amplification.

Error-Correcting Devices. Error-correcting devices are probably the most familiar of the three components of a servomechanism. Commonplace items such as motors, engines, and pistons are all included in the examples of Table 3.

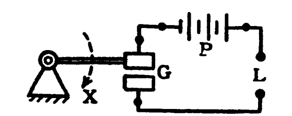
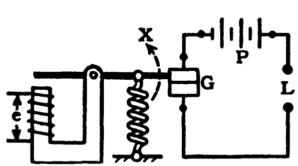
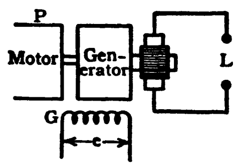
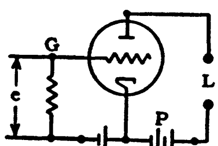
Every error-correcting device as used in a servomechanism has its power input, L, controlled by amplifier, and it in turn uses that power to correct the regulated quantity. This correction of the regulated quantity is indicated by C.

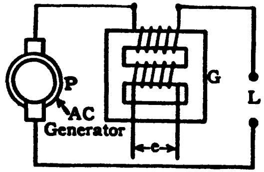
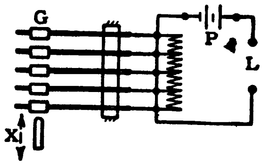
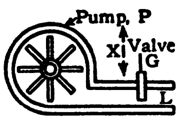
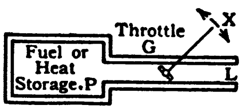
Many present-day servomechanism applications are the direct outgrowth of replacing man as an error detector. The automatic voltage regulator is an excellent example. With manual voltage regulation, the operator uses a meter to measure the regulated quantity, voltage. He detects the difference in needle position from a fixed reference position and adjusts the field rheostat to realign the needle with the reference. This procedure is duplicated by the automatic voltage regulator. It balances the force of a coil, which is a measure of line voltage against the force of a reference spring. When these forces are unbalanced an error in line voltage is detected. The movement resulting from the unbalanced forces produces a change in field excitation that restores the line voltage to its proper value.

Similar comparisons can be made in the case of automatic speed governing, liquid level control, temperature regulation and a multitude of others. In all of these applications the servomechanism does a more accurate and faster job than man. The human transient response is poor when compared to a well-designed servomechanism.

Designing a Servomechanism. The design of servomechanisms is usually broken up into the same three basic categories as given in the schematic diagram, Fig. 3. The designer must select (1) an error detector, (2) an amplifier, and (3) an error corrector that function compatibly with each other and perform the required job. Aside from the usual cost and reliability factors, accuracy and operating force or power determine the selection of the error-detecting device. The amplifier is chosen on the basis of required amplification and time delay introduced. The error-corrector device is determined by the regulated quantity and ease of coupling to the selected amplifier. Many

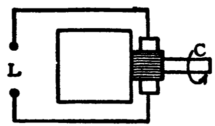
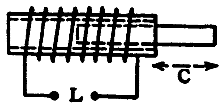
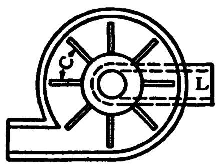
TABLE 2. POWER AMPLIFIERS *

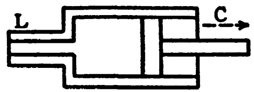
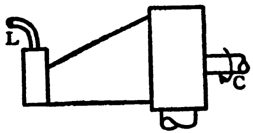
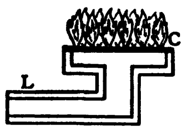
No.	Schematic Representation	Type	"Gate" Element	Approximate Input Units	Approximate Output Units	Approximate Power Amplification Factor	Devices Represented by Load, L
1		Contact	Contact	Ounces	Watts	$1 \times 10^7 \times t$	Relay motor Generator field Impedance Solenoid
2		Relay	Contact	Watts	Watts or kilowatts	1×10^3	Relay motor Generator field Impedance Solenoid
3		Generator	Field	Watts	Watts or kilowatts	50	Motor Impedance
4		Electronic tube	Grid	Microwatts	Watts	1×10^5	Relay motor Generator field Impedance

5		Saturable reactor	D-c coil	Milliwatts	Watts	3×10^2	Generator field Impedance
6		Silverstat	Contacts	Grams	Watts	$1 \times 10^7 \times t$	Generator field Impedance
7		Valve	Valve gate	Inch-pound	Horsepower	$1 \times 10^7 \times t$	Turret Press Heat absorption
8		Throttle	Throttle valve	Inch-pound	Horsepower	$1 \times 10^7 \times t$	Propeller Vehicle Generator Mill

* This table is a condensation of material from the article on servomechanisms by Dr. S. N. Herwald, taken by permission from the September 1946 issue of *The Westinghouse Engineer*

TABLE 3. ERROR CORRECTORS *

No.	Schematic Representation	Type	Input Energy	Output Energy	Approximate Output Power Range
1		Electric motor	Electrical	Mechanical rotation	1×10^{-2} to 4×10^{4} hp
2		Solenoid	Electrical	Mechanical translation	1×10^{-3} to 15 hp
3		Hydraulic motor	Hydraulic	Mechanical rotation	1×10^{-2} to 11×10^{4} hp

4		Piston	Hydraulic	Mechanical translation	1×10^{-3} to 1×10^{3} hp
5		Steam or gas prime mover	Heat or chemical (fuel)	Mechanical rotation	5 to 1.65×10^{5} hp
6		Burner	Chemical (fuel)	Heat	1×10^{2} to 1.5×10^{8} Btu per hr

* This table is a condensation of material from the article on servomechanisms by Dr. S. N. Herwald, taken by permission from the September 1946 issue of *The Westinghouse Engineer*.

combinations of error detectors, amplifiers, and correctors can be assembled to fill any servomechanism job. Often there is little to choose between them.

The use of Tables 1, 2, and 3 illustrates how a servomechanism can be devised. Suppose it was desired to regulate angular position. Where relatively low accuracy is desired, an a-c or d-c resistance bridge, a relay and an electric motor (or a mechanical differential, a contact, and an electric motor) would perform the job. Higher accuracy can be obtained by using an a-c synchro-system with an electronic tube, a generator, and an electric motor; or by using an a-c synchro-system with a valve and hydraulic motor. These components can be chosen from the tables. Each of these is a complete servomechanism and each has its proper place in meeting desired performance requirements. The power of the error-correcting device is determined by the rate at which it is desired to vary the regulated quantity. This rate of varying the regulated quantity also determines the transient errors and the tendency of the system to "hunt."

"Hunting" or self-induced oscillation of the regulated quantity without change in the reference quantity is a problem that has to be faced in all quick-response, high-accuracy servomechanisms. Any closed-loop system such as shown in Fig. 3, where the output feeds back into the input, has a tendency to oscillate.

However, the inherent damping in the simpler servomechanism is usually great enough to overcome this tendency and the system is stable. Where accuracy and response requirements are not too severe, a servomechanism of this self-damped type will perform satisfactorily. To meet the more stringent accuracy requirements, special anti-hunt circuits to give an increased amount of damping are used.

Lately, much attention has been focused on the question of servomechanism stability. The particularly high performance requirements of wartime servomechanism applications have brought this about. The need for the best possible transient response in addition to satisfactory stability requires careful scrutiny of the effects of each type of "anti-hunt" circuit.

In general, there are three approaches to the problem. The first involves getting solutions to the overall differential equation of the system. Since the equations are usually of a high order, the procedure to find the roots is quite difficult. At best, one particular solution can be obtained, and if the effect of varying one parameter is desired the entire root-finding procedure must be gone through again.

The second method is to get the amplitude and phase characteristics of the servomechanism either by direct experimentation or by a variation of the Nyquist method used in feedback amplifier design. It is relatively easy to interpret the phase-amplitude diagram in terms of stability and response. Changing parameters usually requires a new diagram. The third system is to substitute an equivalent electrical circuit for the servomechanism and, by using the transients analyzer, obtain the system response on an oscilloscope as different circuit elements are varied.

Detailed performance and stability calculations are not necessary to understand servomechanism operation. All that is required is a good understanding of Fig. 3. The interrelations shown of the reference quantity, error-detecting device, amplifier, error-corrector device, and regulated quantity provide the basis of all servomechanisms. From combinations of items in Tables 1, 2, and 3 a large number of servomechanisms can be devised.

Examples of Servomechanisms. Three examples of servomechanisms are illustrated in Figs. 4, 5, and 6. Figure 4 shows the components of a heat control in a home. The reference quantity is temperature and the error-detecting device is a bimetallic thermostat located in the space where the temperature is to be controlled. When the space temperature falls below the desired standard, the thermostat closes a low-voltage circuit which includes a relay. When the relay is actuated it closes a second circuit to a motor or to a powerful electromagnet (amplifier). In a coal-burning furnace the motor feeds coal to the burner and furnishes a blast of air for rapid combustion. In an oil furnace a motor pumps fuel oil and air to the burner. For a gas-burning furnace an electromagnetic valve opens and admits gas to the burner. In each of these furnaces the heat produced by the burner is transferred from the furnace to the space above (either by gravity or air-blower action). Thus, the furnace is an error-corrector device which feeds its heat output to the space above and maintains a nearly constant temperature. When this temperature rises above the reference temperature the thermostat (error detector) opens its circuit and stops the motor and the error-corrector device.

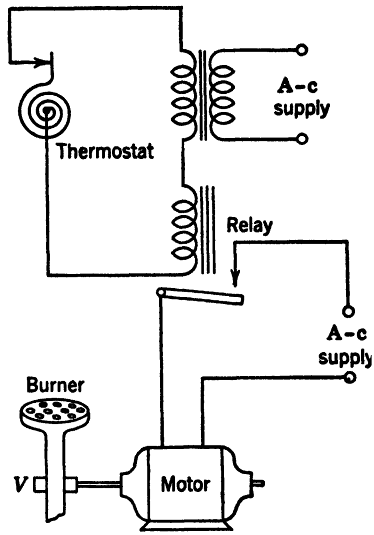


FIG. 4. Furnace-control servomechanism.

A servomechanism for maintaining a constant motor speed is shown in Fig. 5. Here the speed of the motor M is controlled by a variation of the voltage applied to its armature by the generator G. The field of the motor is excited by a separate and nearly constant source of direct current, while the field of the generator G is supplied by a multiple-field exciter (A) such as an amplidyne, a Rototrol, or a Regulex.

The exciter uses a standard reference field as shown and a control field excited by a d-c tachometer (generator) on the motor shaft.

The control field and the standard reference field on the exciter are balanced as long as the motor speed is correct. Any deviation from this speed varies the tachometer voltage and the resultant field of the exciter in such a way as to correct for the change of speed. In this

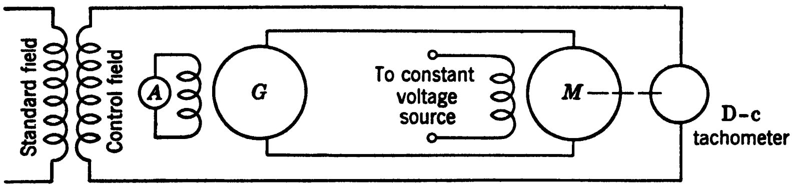


Fig. 5. Motor-speed-control servomechanism using a multifield exciter.

system the reference quantity is motor speed, the error-detecting device is the tachometer, the amplifier is the exciter, and the error-corrector device is the generator.

The feedback amplifier is an important application of the servomechanism principle. A block diagram of the feedback amplifier is shown in Fig. 6. An input signal coming from the left is fed into the amplifier which increases the magnitude of the signal many times. The increase in magnitude of the signal from the input to the output is known as the gain. When amplifiers are used as repeaters in long-distance communication circuits it is very important that the repeater

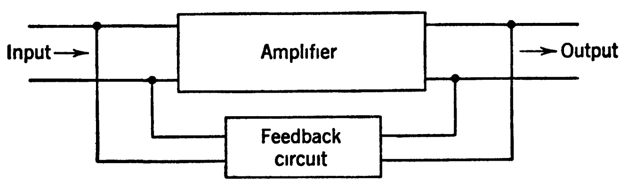


Fig. 6. Negative feedback amplifier.

gain be held constant. Such control may be attained manually but the most satisfactory method is to use the feedback amplifier circuit under discussion. In this feedback circuit a portion of the amplified output is fed back to the input circuit so that it is 180 degrees out of phase with the input. This phase relation causes the feedback to subtract from the signal coming from the line (left). When temperature or voltage changes within the amplifier circuit cause an increase in gain in the amplifier (left to right), an increased signal is fed back through

the feedback amplifier which subtracts a greater value from the incoming signal and thus tends to hold the gain constant. Any decrease in the gain within the amplifier produces the opposite effect with the result that the actual gain is held within close limits at all times. The actual feedback circuit is much more complicated than suggested in Fig. 6. Many other forms of servomechanisms will be explained in the remaining chapters of this book.

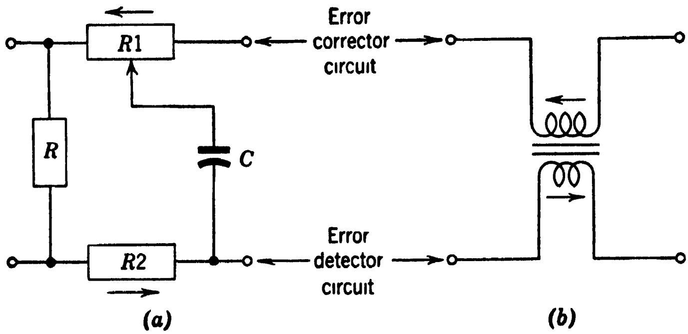


FIG. 7. Antihunt circuits for servomechanisms.

Antihunt Circuits. Antihunt circuits are often required in servomechanisms to prevent "overshooting" or hunting between the error signal and the regulated quantity. Antihunt circuits are of the feedback type and thus are related to the feedback amplifier, but they function in a different manner. In the antihunt circuit the feedback is negative and *proportional to the rate of change* (rather than change only). This feedback functions from the error-corrector device to the error-detector device. The feedback coupling may be produced by an electric field through a capacitor, as in Fig. 7*a*, or through the magnetic coupling of a transformer, as in Fig. 7*b*. Since the feedback is negative and proportional to the rate of change, it has a damping action on all changes in the system and thus prevents hunting.

Chapter XII

RECTIFICATION AND INVERSION

Rectification of electric current is the process of changing an alternating current into a unidirectional current. A unidirectional rectified current is pulsating in character and it is generally necessary to smooth out the pulses by means of a filter in order to utilize the resultant current and voltage. Rectification of alternating current is exceedingly important because (1) it is more simple and economical to generate and distribute electric energy as alternating current, and (2) many applications of electric energy in the field of electric power and communication require the use of a unidirectional or direct current. Some of these applications requiring the use of direct current are the power supply to the anodes of electron tubes, electroplating, chemical processes, charging of storage batteries, and operating series railway motors and adjustable-speed d-c motors.

Rectification of alternating current or the conversion of alternating current to direct current can be brought about by (1) electronic methods and (2) commutating methods. This textbook is concerned primarily with the electronic methods, but for purposes of comparison some of the commutating methods will be given brief consideration. Inversion, the process of converting a direct current to an alternating current, may likewise be performed by both electronic and commutating means.

Commutating Rectifiers. Commutating rectifiers may be of the vibrating contact, the synchronous motor-driven contact, or the rotary type of commutator. A *vibrating rectifier* consists of an electromagnet which vibrates a spring containing movable contacts. The period of vibration is controlled or tuned by the pulses or alternations of the current, and the movable contacts serve to reverse the direction of current flow so that the outgoing current is rectified or unidirectional. Devices employing this principle were used in earlier years for charging storage batteries and for rectification in telephone-ringing machines. Today the commutating vibrator is widely used as

an inverter for changing direct current to alternating current in automobile and other portable radio sets, and in some sets it performs the double function of inversion followed by rectification.

The circuit for a typical vibrating type of inverter is shown in Fig. 1. The closing of switch S energizes the driving coil over a path from the grounded + side of the battery, the driving coil, and the right half of primary winding P to the negative side of the battery. The resulting current energizes the driving magnet and pulls the vibrator V to the

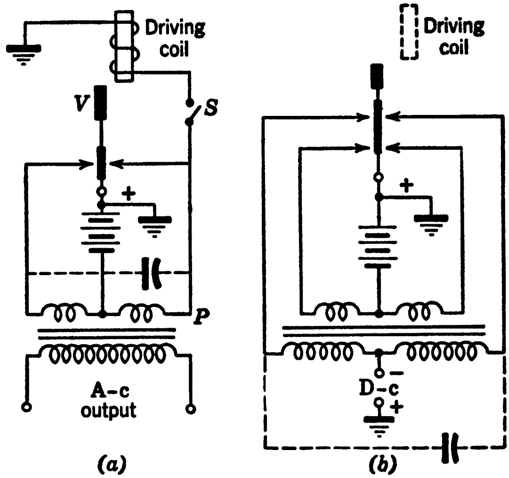


FIG. 1. Vibrator types of inverters and rectifiers: (a) inversion; (b) combined inversion and rectification.

right-hand contact. The closing of this contact (1) sends a relatively large d-c pulse of current through the right half of winding P and (2) places a short circuit around the driving coil. The first action induces a voltage in the output side of the transformer, and the second serves to de-energize the driving magnet. The spring of the vibrator now swings the movable member V to the left until it engages the left-hand contact where it sends a pulse of current through the left half of the primary winding P. This new current pulse reverses the flux in the transformer core and reverses the voltage induced in the output side of the transformer. In the next instant the vibrator V swings again to the right owing to its spring action plus the reestablished current through the driving coil. The repetition of the foregoing action produces an alternating voltage at the output terminals of the step-up transformer. This a-c output voltage may be rectified by a vacuum diode, by a small metallic rectifier, or by the addition of two contacts at the

vibrator and a mid-tapped secondary to the inverter circuit. The additions are shown in part *b* of Fig. 1. There the driving circuit (not shown) acts in the same manner as before. A study of this circuit will show that the added contacts serve to reverse the voltage between the mid-tap of the output winding of the transformer and ground in such a way as to make the resulting voltage pulses unidirectional, and thus give full-wave rectification. A smoothing filter must be added to give an output suitable for the power supply for radio receiving sets.

The preceding explanation has been an over-simplification of the complete theory of action involved in the vibrator type of inverter and combined inverter and rectifier. Condensers are used in the circuits to reduce arcing at the contacts but the complete theory of action and design involves a proper balance and tuning of the L and C constants in the circuits.*

A new type of mechanical rectifier or *contact converter* was developed by Siemens-Schuckert in Berlin during World War II. This contact converter is capable of rectifying currents as high as 10,000 amperes at 400 volts with efficiencies of the order of 98 to 99 per cent. The principle of rectifying alternating current by reversing the circuit in step with the alternations is old in the art, but for sizable currents no satisfactory process has been developed until recently. In theory it is necessary only to reverse the circuit while the voltage and current pass through the zero point on the cycle. To effect this requirement has been impractical because of the short time available for mechanically moving the contacts at the time when the rate of current change is the maximum. Any phase-angle difference between the voltage and the current naturally complicated the problem of commutation. If the wave form of the circuit can be modified to hold the voltage at near zero for a short time interval at the points of voltage and current reversal, it is obvious that the mechanical problem of reversal would be simplified. This important advantage is brought about in the new contact rectifier by inserting saturable inductances or chokes in the alternating-current leads. These chokes saturate at relatively low values of flux and current, so that they produce a high impedance to current change at low values of current (region when passing through zero) and little impedance at higher current values. These chokes act in a manner similar to the peaking transformer although their function is much different. The application of the saturable choke gives the necessary delay in the voltage and current change and permits the mechanical reversal of contacts operated by a small

* For more complete details the reader is referred to the *Vibrator Data Book* published by P. R. Mallory & Co., Inc., copyright 1947.

synchronous motor. Many mechanical and electrical refinements are necessary to make the contact converter successful. The discussion of these refinements is out of the province of this book. Contact converters are now being manufactured in the United States.*

Rotary commutating rectifiers have been of three types: simple commutators, motor-generator sets, and rotary converters. The simple commutator device consists of a synchronous motor driving a commutator at such a speed that the polarity is reversed at the proper rate to give a rectified or unidirectional current in the output. This device was widely used in early radio transmitters and X-ray equipment and today has considerable application for rectifying high voltages for precipitation processes in the removal of smoke, dust, and chemical byproducts. The motor-generator rectifier consists of an a-c motor operating from an a-c system which drives a d-c generator. This process is an electrical-mechanical plus a mechanical-electrical conversion process. It has the advantage of good control of the d-c output voltage and the disadvantages of high initial cost, lower efficiency, and the complications of rotating equipment. This system is being replaced by electronic methods of conversion. The rotary converter consists of a single rotating unit which consumes alternating current in motor action and gives forth a d-c output. The output is partly the result of pure rectification and partly of a conversion process. This device has a lower cost and a much higher efficiency than the motor-generator device. Its disadvantage lies in inflexibility in the control of the output (d-c) voltage. This unit once had a wide application in the electric railway field but is being replaced by electronic rectifiers.

All these commutating devices used for conversion of alternating current to direct current may be employed equally well through a reverse process to bring about inversion.

Electrolytic Rectifiers. An electrolytic rectifier consists of two dissimilar electrodes placed in an electrolyte. One type of cell uses an aluminum plate and a lead plate for the electrodes and a solution of ammonium phosphate for the electrolyte. In this cell electrons pass readily from the aluminum electrode to the lead electrode but not vice versa. Thus the cell may be used as a single-wave rectifier. Obviously, the phenomenon is of electronic origin and the action is explained by the presence of a barrier at the surface of one electrode which has a high resistance to the passage of electrons from the electrolyte into the heart of the electrode but little resistance to a movement in the reverse

* For more information the reader is referred to the article, "A Mechanical Rectifier" in the April 1948 issue of *Power Generation*.

direction. The work function of the material at the surface of the electrodes determines the resultant action.

The electrolytic rectifier was widely used for charging storage batteries during the pioneering days of radio. It had the advantage of simplicity and low cost and the disadvantage of low efficiency and short life. Its low efficiency was due to a rather high internal voltage drop plus a loss from a small reverse current flow. The electrolytic rectifier has been obsolete since 1930.

Single-Phase Rectifier Circuits. The simplest circuit for rectifying single-phase alternating current gives half-wave rectification. Such a circuit is shown in Fig. 2, part *a*. This circuit can be simplified by reducing it to the schematic or equivalent form of part *b* in the same figure. Here a new symbol represents the rectifying unit and *the arrow of the symbol indicates the direction of conventional current flow.* When considering this or any other rectifying circuit the student may wish to know the magnitude of the d-c voltage and current, the regulation of the load voltage, and the efficiency to be expected from the rectifying process. All these values depend upon a number of variables such as the transformer constants and characteristics, the characteristics of the rectifying unit, and the nature of the load. These variables may make an analytical solution of the problem rather difficult. However, it is possible to make certain assumptions which reduce the circuit to an ideal basis from which a helpful analysis may be made. Referring to Fig. 2, part *a*, assume (1) an ideal transformer without resistance, without core loss, and with zero regulation, (2) that the impressed emf is a pure sine wave, (3) that the rectifying unit has zero resistance in the forward direction of electron movement and infinite resistance in the reverse direction, and (4) that the load is a pure ohmic resistance. With these assumptions, let an alternating voltage of effective value E_{rms} be impressed on the left side of Fig. 2, part *b*. The impressed voltage E_{rms} may be represented by the sine wave of part *c* of Fig. 2 and this voltage will produce a rectified half-wave of current as shown in part *d* of the same figure. This half-wave of current is of a sine form since $i = (E_m/R_L)\sin\omega t$. During the last half of the cycle the rectifier will block the electrons and no current will result. The same loop or half-wave of current will flow on both the input and output sides of the rectifier. This current flowing through the load resistance R_L will produce an iR_L voltage drop which is likewise of a sine form. The d-c values of pulsating wave forms are the average values. By the methods of calculus, the area of the rectified sine wave form of current in Fig. 2*d* is the integral from 0 to π of $i\,d\theta$

which equals $2I_{max}$. The average value of this area spread over a complete cycle is

$$I_{avg} = I_{dc} = \frac{2I_m}{2\pi} = 0.318I_{max} \tag{1}$$

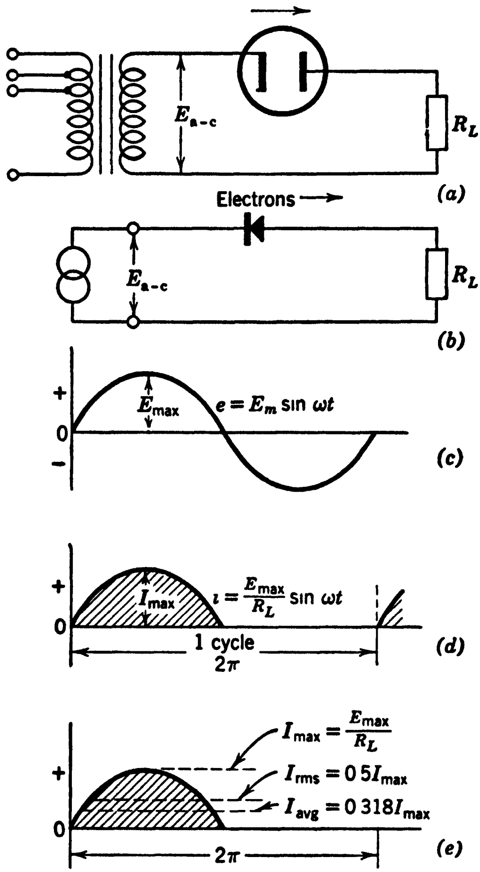


FIG. 2. Circuit and waves for half-wave rectification.

The instantaneous voltage equals iR_L, and the same deductive process shows the average voltage on the load side of the filter for the complete cycle to be

$$E_{avg} = E_{dc} = 0.318E_{max} = 0.45E_{rms} \tag{2}$$

The effective value of the current on the input side is the root-mean-square value. Obviously, the effective value of the first half-cycle of current is $E_{max}/\sqrt{2}$. The mean square of this is $E_{max}{}^2/2$. Now if this

latter value is spread over the complete rectifying cycle, its mean square value is

$$\frac{1}{2} \times \frac{I_{\max}{}^2}{2} = \frac{I_m{}^2}{4}$$

and the root-mean-square of the latter is

$$\frac{I_{\max}}{2} = 0.5I_{\max} \tag{3}$$

These values are indicated in part *e* of Fig. 1. Summarizing,

Input Side	Output Side
E_{rms} = effective impressed voltage	$E_{\text{dc}} = 0.45E_{\text{rms}}$
$I_{\text{rms}} = 0.5I_{\max}$	$I_{\text{dc}} = 0.318I_{\max}$

The heating losses (I^2R) which occur in the various parts of the complete rectifier circuit are increased by the irregular wave forms of current which are inherent in the rectifying process. This may be illustrated by comparing the heating loss produced in the load resistance R_L by a smooth or filtered d-c current with that resulting from the current wave of Fig. 2. The resulting ratio may be termed the ratio of rectification. Thus from the preceding table

$$\underset{\text{(half-wave rectification)}}{\text{Ratio of rectification}} = \frac{(I_{\text{dc}})^2R_L}{(I_{\text{rms}})^2R_L} \times 100 = \frac{(0.318I_m)^2}{(0.5I_m)^2} \times 100$$

$$= 40.6\% \tag{4}$$

The ratio of equation 4 is sometimes called the efficiency of rectification. The latter terminology is somewhat misleading since the overall power efficiency for the assumed conditions (zero losses) must be 100 per cent. The real significance of the ratio of rectification is that it gives a qualitative indication of the increased heat losses that occur wherever a nonsinusoidal and pulsating current flows through resistance elements in the actual rectifier circuit such as the transformer windings and in the load without a filter. A second method of expressing the increased heat losses is by the term current *form factor* which is the ratio of root-mean-square to average value.

The capacity of the input transformer to a rectifier is used inefficiently when supplying power for half-wave rectification because only one-half of the sine wave of current is passed and the secondary winding carries a d-c component of current which magnetizes the iron core and increases the core losses. These factors plus the increased I^2R losses suggested in the preceding paragraph result in a reduction of the

permissible transformer output under rectifier loads. This reduction is sometimes expressed by the term utility factor. *Utility factor* is

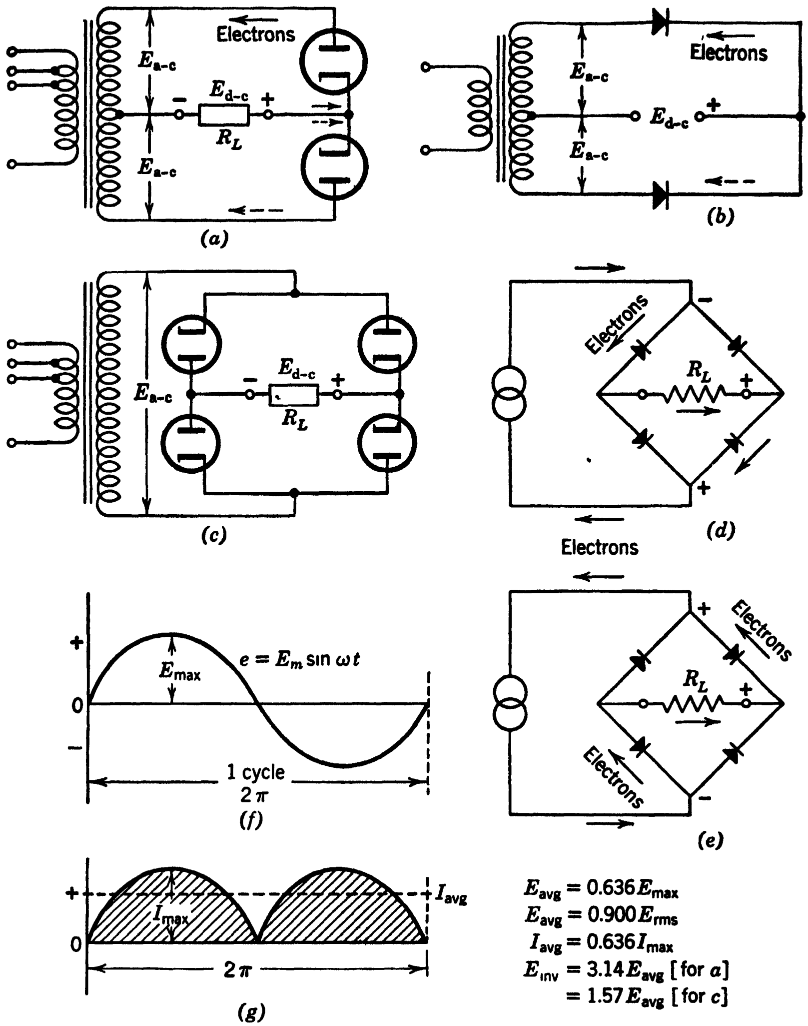


FIG. 3. Circuit and waves for full-wave rectification. (Conventional current flows in direction of boldface arrows.)

the ratio of the permissible rectifier load without overheating to the normal unity power factor load that the transformer will carry.

There are two types of circuits used for full-wave single-phase rectification. One of these circuits using a transformer with a mid-tap in the secondary winding is shown in parts *a* and *b* of Fig. 3. For one

half-cycle of the impressed a-c voltage E_{rms} the upper half of the transformer secondary and the upper diode conducts current, and for the second half-cycle the lower half of the transformer and the lower diode operates. If a pure sine wave of voltage like part *f* of Fig. 3 appears across each half of the transformer secondary winding, both halves of the cycle will be rectified as shown in part *g* of this figure. The connection to the midpoint on the secondary of the transformer serves to reverse the polarity on the lower rectifier with respect to the upper one. Since both half-waves of current pass through the transformer (though dividing in the secondary), there is no d-c component of flux in the transformer core to increase core losses, and the current in the primary is normal. The second type of circuit to give full-wave rectification is shown in parts *c, d,* and *e* of Fig. 3. This is known as a bridge circuit since it uses four rectifying units connected in a form similar to the Wheatstone bridge. The rectified currents flow through two diodes in series and the action for the two halves of each cycle is clearly shown in parts *d* and *e* of the figure. The resulting impressed voltage and current waves are the same as for the first full-wave rectifier circuit (Fig. 3, parts *f* and *g*).

For the same assumptions made for the preceding half-wave rectifier circuit, the relation between the input and output sides of either full-wave rectifier circuit may be readily calculated. Since both halves of the cycle are rectified, the current and voltage on the input side are normal effective values and those on the d-c or output side are average values. Thus

$$E_{rms} = \text{applied rms value} = \frac{E_{max}}{\sqrt{2}}$$

$$I_{rms} = \frac{E_{rms}}{R_L} = \frac{E_{max}}{\sqrt{2}R_L} = 0.707\frac{E_{max}}{R_L}$$

$$E_{dc} = E_{avg} = 0.636E_{max} = 0.9E_{rms}$$

$$I_{dc} = I_{avg} = 0.636I_{max} = 0.9I_{rms}$$

$$\underset{\text{(full-wave rectification)}}{\text{Ratio of rectification}} = \left(\frac{I_{dc}^2R_L}{I_{rms}^2R_L}\right) \times 100$$

$$= \left(\frac{0.636}{0.707}\right)^2 \times 100 = 81.2\% \qquad (5)$$

Equation 5 gives a concept of the extra I^2R losses occasioned by the wave form of the resulting currents.

The rectified voltage and current output of any rectifier consists of a series of unidirectional waves or ripples. For some applications these variations are not objectionable but for others they must be smoothed out by filters. For all cases the relative magnitude of the ripple is important in the comparison of rectifying circuits. The comparison is made in terms of ripple factor. *Ripple factor is the ratio of effective value of the alternating components of the rectified voltage or current to the average value.* The alternating component of current for half-wave and full-wave rectification using a sine wave is shown by the shaded portions on Fig. 4. The d-c or average component is represented by all the area underneath the dotted line for I_{avg}. The ripple factor for the half-wave rectification is 1.21 and for full-wave rectification is 0.48. The d-c and a-c components of rectified current can be measured by inserting a d-c and an a-c ammeter (plus current transformer) in series in the circuit.

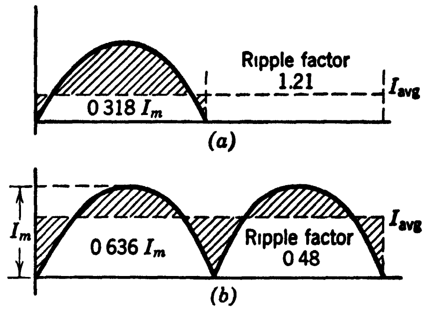


FIG. 4. Alternating component of rectified current: (a) half-wave; (b) full-wave.

The preceding discussion of ripple factor and a further inspection of Fig. 4 will explain why the ratio of rectification is lower than might be expected. The rectified current may be considered as consisting of two component currents, one a steady direct current of the average value useful for d-c application and the other an a-c ripple current that contributes to the I^2R or heat loss in the load. This heat loss may constitute a sizable factor in reducing the efficiency of d-c motor loads operating on rectifiers. In half-wave rectification the a-c component is very large (1.21 times) compared to the d-c component.

A further comparison of the three single-phase rectifying circuits may be of interest. The half-wave rectifier circuit is simple, uses only one rectifier unit, has a high ripple factor, makes inefficient use of its transformers, and has a low efficiency of rectification. The full-wave rectifier circuits have a smaller ripple factor and a much higher ratio of rectification. The circuit using a mid-tap transformer requires only two rectifying units but the transformer is special and more costly. Besides the necessity of a mid-tap, the secondary must have twice as many turns of somewhat smaller wire; hence a higher cost. In addi-

tion, the distorted wave form of the current in the two secondaries adds to the I^2R losses. Where diode rectifying tubes are used this circuit is standard because tubes have a relatively high initial cost and maintenance cost. The full-wave bridge circuit uses a simple two-circuit transformer but four rectifying units. This circuit is standard for use with the blocking-layer rectifiers because the rectifier cells required in series for a given voltage can be separated into two parts to comprise the two rectifying units for each path through the bridge circuit. This circuit uses a simple transformer with a single winding on the secondary. Under the ideal conditions originally assumed, there are no extra losses in this transformer arising from rectification. Under actual conditions of rectification there is some wave distortion resulting in a utility factor of less than one.

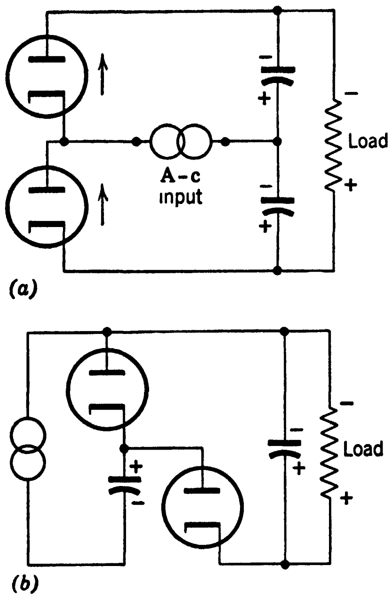


FIG. 5. Voltage doubler circuits: (a) full-wave; (b) half-wave.

The rectified output voltages of the single-phase circuits considered above are less than the input effective voltages. If an output voltage greater than the input is desired, it can be obtained by either of the voltage doubler circuits given in Fig. 5. In part *a* full-wave rectification is secured by connecting two diodes in two arms of a bridge circuit with condensers in the other arms. While rectifying one-half a cycle of an a-c voltage, the upper diode will charge the upper condenser to a potential approaching E_{max}. Similarly, on the second half-cycle, the lower diode will charge the lower condenser to a potential approaching E_{max}. Since the diodes rectify voltages that are series aiding, the total voltage across the load and the two condensers in series will approach the value $2E_{max}$. The actual voltage will depend upon the current drain (discharge) by the load and upon the cathode-anode voltage drop across the diodes. Part *b* of Fig. 5 shows another circuit for doubling the output voltage. Here on the first half-cycle of impressed voltage the diode on the left charges the condenser on the left to a potential approaching E_{max}. On the second half-cycle the peak voltage

across the diode to the right is the impressed E_{max} plus the voltage due to the charge on the left-hand condenser. The rectification of this combined voltage approaching $2E_{max}$ charges the condenser on the right to a load voltage approaching $2E_{max}$ as in the first circuit. The circuit of part *a* gives better regulation, higher ripple frequency, and lower voltage across the condensers; the circuit of part *b* has a common input and output terminal.

The actual resistance and power loss of a rectifying unit will alter the voltage and current relations and efficiencies developed in the preceding discussion where an ideal status was assumed. The rectifying units in any of the circuits considered may be high-vacuum diodes, gas or vapor diodes, or blocking-layer rectifiers. The differences in the characteristics of each of these units were discussed in earlier articles. It is possible to make some generalization covering these characteristics and thereby make a closer approximation for some factors of rectification. Thus for vacuum diodes and blocking-layer units the assumption may be made that the resistance is constant. If the resistances of such a rectifier are represented by the symbol R_R, leaving the load resistance R_L, then for any given set of conditions the rectified current and the rectified voltage will be reduced in the following ratio.

$$\underset{\text{(high-vacuum tubes)}}{\text{Ratio of reduction}} = \frac{R_L}{R_L + R_R} = \frac{1}{1 + (R_R/R_L)} \tag{6}$$

In gaseous and vapor tubes the arc drop or cathode-anode drop is practically constant. If this drop is represented by V_R, then the instantaneous values of the output voltage will be the impressed value minus V_R as shown in Fig. 6. Any general analytical solution for this case becomes somewhat involved and of doubtful value.

The output voltage E_{dc} of a rectifier using vacuum diodes or blocking-type cells will decrease with load by an amount determined by the reduction factor (equation 6). If it is necessary to hold the value of E_{dc} to a constant value, it is necessary to vary the input voltage E_{rms}. This can be accomplished in steps by changing taps on the primary of the input transformers in Figs. 2 or 3, or more exactly by some form of input voltage regulator.

Multiphase Rectifier Circuits. Multiphase rectifier circuits are used whenever moderate or large magnitudes of d-c power are applied. Such circuits have the advantage of utilizing the standard three-phase power distribution system and they furnish smoother rectified voltage and current waves at higher efficiencies. A simple three-phase circuit illus-

trating the principle of multiphase rectification is given in Fig. 7. Power is supplied through a three-phase, delta-wye connected transformer to a three-anode unit rectifier. The rectifying units may consist of vacuum diodes, gaseous diodes, phanotrons, ignitrons, or excitrons having all cathodes connected together, or a multielectrode tank rectifier having a single-cathode mercury pool may be used. The d-c load is connected to the common cathode terminal (+) and to the neutral of the transformer secondary. All half-wave multiphase circuits require that one load terminal arise from a neutral or common point in the transformer supply secondary. The effective a-c voltage for rectification E_{rms} is always the voltage measured from the transformer neutral to the anodes of the rectifier.

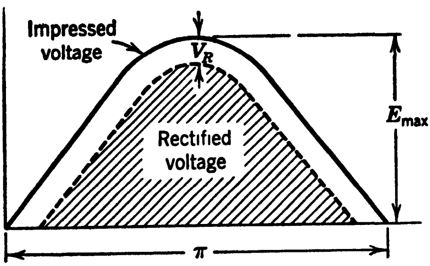


Fig. 6. Effect of a constant arc drop upon the output voltage of a rectifier.

An analysis of the action of multiphase rectification may be made by assuming ideal rectifying units with zero resistance, an ideal transformer, a pure resistance load, and a sine wave of applied voltage. The variation of voltage across the anodes during one cycle is depicted in the middle section of Fig. 7 and the resulting current waves in the lower section of the figure. The d-c voltage waves will have the same form as the current waves. The rectified waves have ripples per cycle equal to the number of anodes but the ripples are lower in magnitude giving a lower ripple factor and a much smoother wave. Under the ideal theoretical conditions assumed, the rectified current flows to the anode having the highest positive potential and that means that each of three anodes will carry the current for $2\pi/3$ part of each cycle. Since the ripples on the load side of the transformer are equal and continuous, the d-c or average value of the current or voltage will be the average value of a single ripple wave. The peak value of the ripple voltage is the maximum value of the impressed voltage E_{max}. As the number of phases and anodes increase, the number of ripples and the average d-c voltage increase and the voltage approaches the maximum value of E_{max}.

One method of obtaining a low ripple from a three-phase circuit and a standard transformer is shown in the circuit of Fig. 8. In practice, this circuit, using a delta-connected transformer, generally em-

ploys the blocking-layer rectifier. Six rectifier tubes of any type can be employed though the higher cost of the tubes may make their use uneconomical. The transformer phase voltages and the resulting current output with six ripples per cycle is shown at the right of Fig. 8.

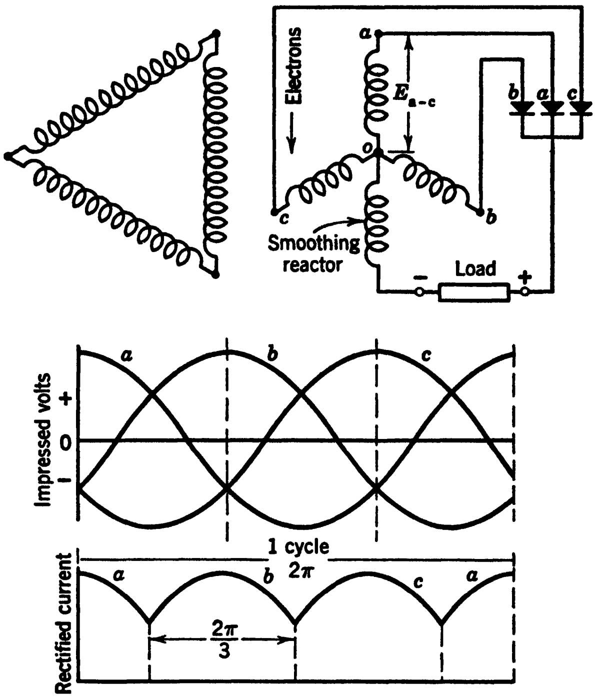


FIG. 7. Half-wave, three-phase rectifier circuit and voltage-current waves (boldface arrows show direction of conventional current).

Obviously, this three-phase circuit is analogous to the full-wave single-phase bridge type of circuit.

A schematic circuit for six-phase rectification using a special transformer and a rectifier having six anodes is given in Fig. 9. This circuit is called the three-phase diametric (six-phase star). Twelve and occasionally eighteen anodes are used for rectification. The ratio of voltage conversion, that is, volts of direct current to volts of alternating current effective for a few multiphase circuits, is given in Table 1.

TABLE 1

ANODES	3	6	12	18
$\frac{E_{dc}}{E_{rms}}$	1.17	1.35	1.40	1.41

The voltage ratios given for the ideal circuit do not hold for the actual circuit because of the voltage regulation of the transformer and the cathode-anode drop in the rectifying unit. Hence the load voltage reg-

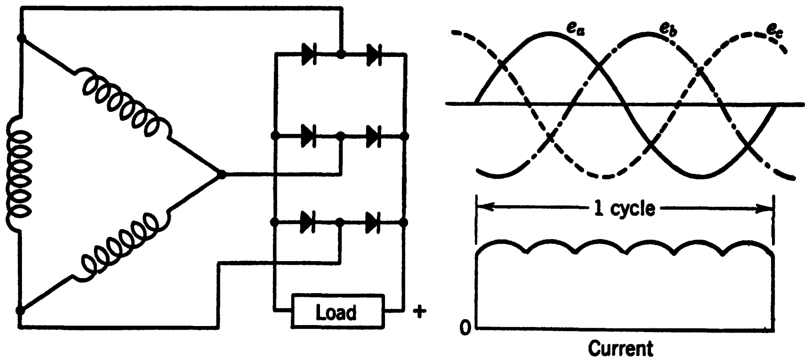


FIG. 8. Full-wave bridge, three-phase rectifier with voltage and current relationship.

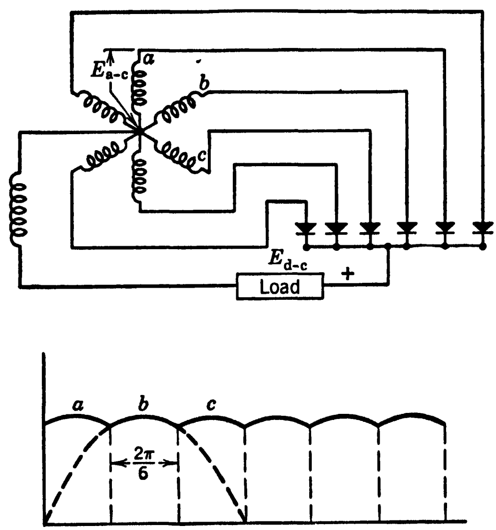


FIG. 9. Schematic six-phase rectifier circuit and wave form.

ulation of the multiphase circuit may be too poor to meet the requirements of an application. Control of the output voltage can be effected in two ways. One method is to control the transformer input voltage by a regulator or by using taps on the primary side of the input transformer. These methods are not well adapted where large amounts

of power are involved. The second method is to control the time of firing of the anodes by grid action. A delay in the firing of each anode will change the wave form of the rectified current and voltage and reduce the magnitude of the average value. The application of this method is suggested in Fig. 10. This method of voltage control increases the ripple and the ripple factor.

In the ideal multiphase rectifier circuit the anode at the highest positive potential carries the entire d-c load. As the potential of the conducting anode falls to that of a succeeding anode, the former drops the load and the latter picks it up so that only one anode is conducting at a time. In the actual rectifier circuit this sudden change of load requiring an infinite rate of change of current for the anodes participating in the action does not take place. The transformer secondary winding that supplies each anode has some leakage reactance which opposes the sudden change of current. Accordingly, the current in the first conducting anode tapers off and the current in the second anode rises as the two anodes pass each other with regard to positive potential. Thus in practice there are periods of transition when two anodes are conducting at the same time. In special cases, such as twelve- or eighteen-anode circuits, there may be three or more anodes conducting simultaneously. A twelve-phase ignitron rectifying unit is shown in Fig. 22.

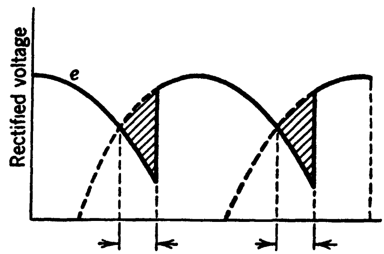


FIG. 10. Voltage control through delay in firing time.

In multiphase rectifier circuits, as in single-phase circuits, the part-time rectification of current from each phase produces distorted (non-sine-wave) currents in the transformer secondaries and may produce residual magnetomotive forces and fluxes in the cores of the transformer feeding the rectifier. These things introduce problems in the design and selection of transformer connections to give optimum results and efficiency. An analysis of these problems lies in the realm of the study of power transformers. However, it should be noted that simple multiple-phase circuits such as shown in Figs. 7 and 9 are seldom employed. The high peak-to-average demand on the rectifying elements together with poor transformer utilization, efficiency, and regulation dictate the use of multiple-conduction circuits using interphase and other special transformer connections like those illustrated

in Fig. 11. Thus a rectifier may sometimes be referred to as a 36- or 72-phase rectifier (common in electrochemical installations), though it is generally made up of a number of three-phase wye groupings properly phased and interconnected with interphase transformers.

The power losses in multiphase rectifiers consist of (1) those in the

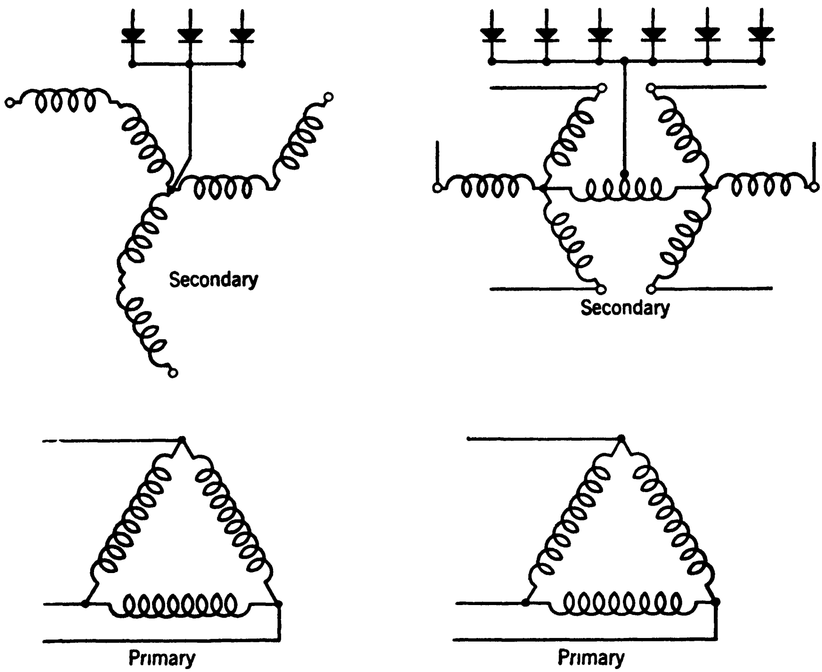


FIG 11 Delta-zigzag and interphase transformer connections for three- and six-phase rectification

input transformer, (2) those in the rectifying unit, plus (3) those in auxiliary equipment. Transformer losses under load are relatively low and the full-load efficiency of transformers is very high. In the rectifying units the losses consist of power for producing thermionic emission (if thermionic cathode is employed) plus the loss in the cathode-anode circuit. The cathode-anode fall of potential is of the order of 10 to 20 volts and if the output voltage is 600 volts or higher, the percentage of power lost across the rectifying unit is very low. Thus for 600 volts and an arc drop of 15 volts the power loss in the arc is 2.5 per cent, and for a load voltage of 3000 volts with 15 volts arc drop the power loss is only 0.5 per cent. The overall efficiency of

complete rectifying units of moderate and large size may well be of the order of 95 to 97 per cent.

Summary on Rectifiers. A summary covering some of the important factors and applications of rectifying units is given in Table 2.

TABLE 2. SUMMARY ON RECTIFIERS

Device	Cathode Content	Load Voltage Range	Load Current Range	Inverse Voltage Peak	Applications
Kenotron (high-vacuum diode)	Hot (vacuum)	0–100,000	Few amperes	100,000	High voltage-low current plate supply for X-ray tubes, radio transformers, Precipitron, etc
Gaseous rectifier (Tungar)	Hot (gas)	0–220	0–15	300	Charging storage batteries.
Phanotron	Hot (vapor)	0–2,500	0–30	5,000	Charging storage batteries. Generator and motor fields.
Thyratron	Hot (gas or vapor)	0–100,000	0–100	100,000	Control equipment. Lighting, timing, relays, heating, etc.
Grid-glow tube	Cold (gas)	0–300	0–0 0001	600	Control for sensitive relays.
Ignitron	Mercury pool	0–3,000	0–∞	3,000	Power rectification, resistance welding, inversion.
Metal-tank rectifier	Mercury pool	0–3,000	0–∞	3,000	Power rectification, inversion.
Excitron	Mercury pool	0–3,000	0–∞	3,000	Power rectification.
Blocking-layer rectifiers	Cold metal plate	7–15 volts per cell	0–1,000	5–50 volts per cell	Battery charging, electroplating, motion-picture arcs, cathodic protection.

Smoothing Filters. A smoothing filter is a circuit element or a network designed to reduce the ripples of current and voltage in the output of a rectifier. Some applications of direct current can utilize the pulsating output of a rectifier without modification but in the majority of cases a smoothing out of the ripples is necessary. The basic elements for smoothing action are condensers and inductance or choke coils. The former act through their ability to store energy in electric fields and the latter through a similar storage of energy in magnetic fields.

The simplest circuit for a half-wave, single-phase rectifier consists of a condenser placed in parallel with the load on the output side of the rectifier as shown in part *a* of Fig. 12. Assuming (1) an impressed voltage having a sine wave, (2) a rectifying unit (low resistance in forward direction), and (3) a high resistance load R_L, the voltage and current values in the circuit will vary as depicted in parts *b* and *c* of Fig. 12. First, the impressed voltage follows dotted line e_s which will charge the condenser C as it rises to the first peak. Then, as the impressed voltage falls below the voltage across the condenser, the rectifier will cut off and the energy stored in the condenser will discharge through the resistance R_L in accordance with the transient equation:

$$i = -\frac{E^{-t/RC}}{R}$$

The voltage across the condenser and load will decrease, as shown by the full line e_R, until the next positive loop of impressed voltage rises above e_R and then the condenser charges again. The current i_R de-

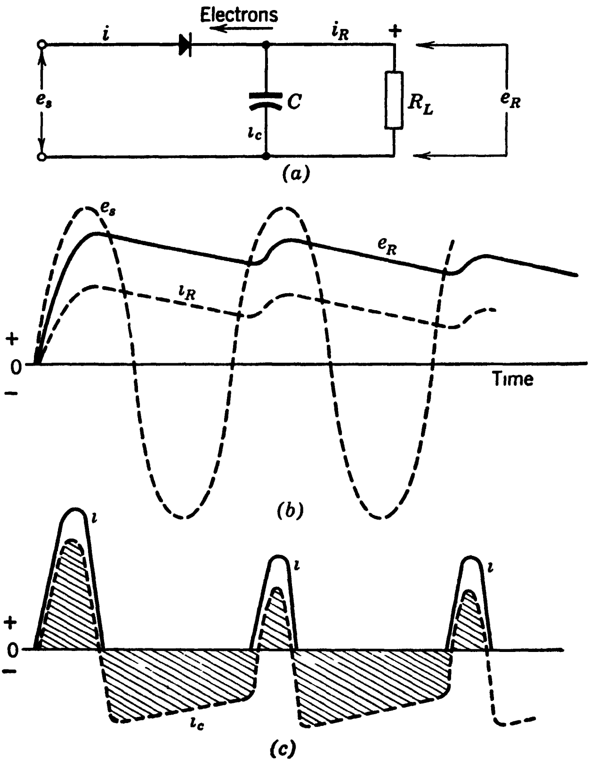


FIG. 12. Half-wave, single-phase rectifier with a simple condenser filter.

livered to the load follows the trend of the voltage across the load. The current i flowing through the rectifying unit rises in pulses as shown by the full line in part c, Fig. 12, and the charge and discharge current for the condenser follows the trend of the dotted line i_c. It should be noted that the pulses of current through the rectifier have a high transient peak and will damage the cathode of a gaseous or vapor-rectifying tube. Obviously, the magnitude of the load current i_R and the magnitude of the voltage and current ripples will be determined by the values of C and R_L. In the actual rectifier the rectifying unit has some resistance which produces a voltage drop and reduces the

voltage across the load. As R_L approaches ∞, e_R will approach E_{max}.

A simple condenser filter connected in a full-wave rectifier circuit as shown in part *a* of Fig. 13 will produce voltage and current waves as illustrated in part *b*. In comparison with the half-wave rectifier of the preceding paragraph, this circuit has output ripples of double the fre-

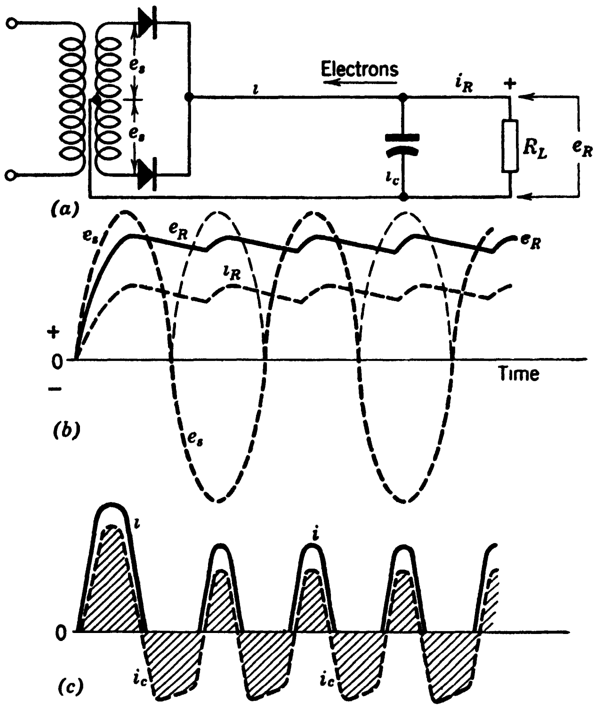


FIG. 13. Full-wave, single-phase rectifier with a simple condenser filter.

quency but of lower ripple magnitude and gives a smoother output. The pulses of input current i and the condenser current i_c are indicated in part *c*.

A simple choke filter connected in the output of a half-wave single-phase rectifier as shown in part *a,* Fig. 14, will produce the current and voltage waves given in part *b*. The inductance of the choke will retard the rise of current through the rectifier and, after the impressed voltage reaches its positive peak, the energy stored in the magnetic field will continue the d-c pulse after the applied voltage becomes negative. Thus the conduction period will continue for more than one half-cycle, though the rectified current will be extinguished and appear as

pulses. This peculiar action is explained by the induced voltage across the choke as illustrated by the curve e_x. The voltage across the load

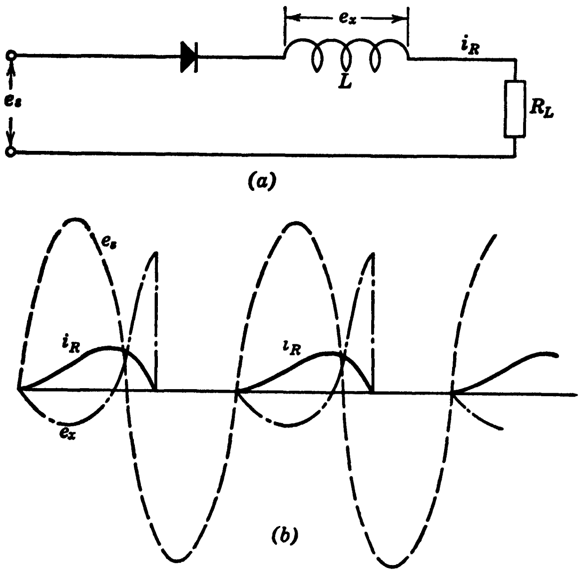


FIG. 14. Half-wave, single-phase rectifier with a choke filter.

will have the same pulse wave shape as the current. In a full-wave rectifier without filter the current falls to zero at the end of each half-cycle. Here a simple choke filter will serve to reduce the peaks and

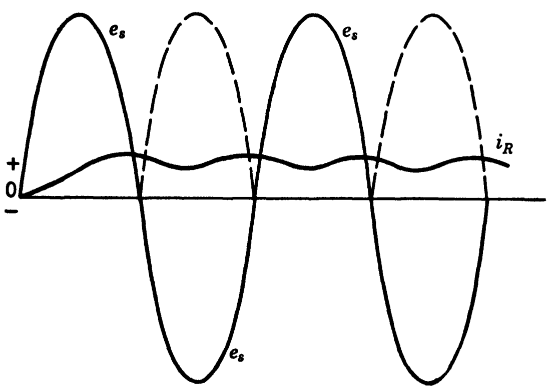


FIG. 15. Full-wave, single-phase rectifier with a choke filter.

prevent the current from falling to zero, as shown in Fig. 15. Likewise, for multiphase rectification where the number of ripples is increased

and the ripple factor reduced, the simple inductance filter will be very effective in smoothing out the ripples. The voltage wave form across a resistance load will follow the trend of the current wave.

Filter circuits using combinations of the condenser and chokes will

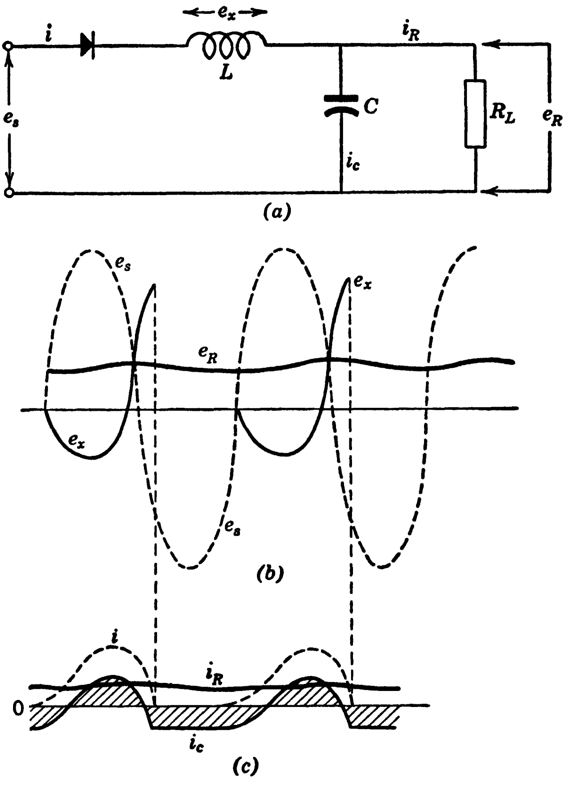


FIG. 16. Voltage and current relations in an L-type filter.

give more effective smoothing action. One simple circuit combination is shown in Fig. 16. Here the series choke will retard the rise and fall of current input to the shunt condenser and load. This action will reduce the magnitude of voltage across the condenser as well as reduce the change in voltage applied to the condenser and load. The condenser, in turn, will absorb energy on the higher voltage inputs and release this energy when the input voltage is low. Both actions tend to maintain a constant voltage across the load and give a good filtering action. The voltage and current relations in the choke input L-type of filter are shown in parts *b* and *c* of Fig. 16. A second form of L filter

uses a shunt condenser on the input side as in Fig. 17. In this circuit the condenser charges rapidly with the rise of rectifier output and to a much higher voltage than in the preceding case. The series choke, in turn, retards the change of current flow to the load and tends to hold constant the voltage at the load. Obviously, the voltage at the load is higher than in the preceding choke input circuit.

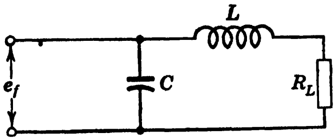


FIG. 17. L-type filter with a condenser input.

The preceding simple filter sections consisting of condensers and inductances in "lumps" are frequently called *brute-force filters.* These filters provide satisfactory filter action for many applications at a minimum cost. Better filtering action will be obtained if the capacitance and inductance are broken into smaller units and assembled in π and T sections, as shown in Fig. 18. The cost of the complete filter having multiple units is greater because the cost per unit of capacity or inductance is increased as the capacity of the unit decreases. Hence economy as well as the excellence of filtering action must determine the selection of the filter design. In applications where the load current is of low magnitude and where cost or weight of apparatus are critical, resistors are sometimes substituted for chokes in smoothing filter circuits.

Comparison of Filter Circuits. The simple condenser and all other condenser input filter circuits cannot be employed with hot-cathode gaseous and vapor rectifier tubes. This follows because (1) the impedance of a condenser to its charging current is very low, and (2) gaseous and vapor rectifier tubes will not withstand cathode-anode potential drops above approximately 28 volts during the conduction period without damage to the cathode arising from positive ion bombardment. Vacuum-tube and blocking-layer rectifiers can be used satisfactorily with condenser input filters.

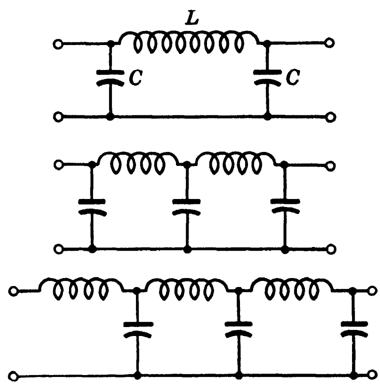


FIG. 18 Combinations of π and T sections to give superior filtering action.

Economy and performance determines the selection of the proper filter circuit. The simple parallel condenser is low in cost for half-wave rectification. The condenser input L filter gives a higher load to input-voltage ratio but is likely to give poorer load-voltage regulation. Since the cost of condensers rises rapidly with voltage rating, the choke input may prove more economical for higher load voltages. The choke input L type gives a lower load to input-voltage ratio but has better inherent voltage regulation. The simple series choke filter is usually satisfactory for filtering with multiphase rectifier circuits.

Precautions on Peak Inverse Voltage. In designing rectifier circuits and the accompanying filters, the engineer must give due consideration to the peak inverse voltage to which the rectifying units will be subjected. As pointed out on page 113, the presence of a condenser filter on a half-wave rectifier may subject the rectifying unit to an inverse peak approaching two times the a-c maximum voltage. Transient voltage surges on supply lines or transients due to transformer switching may increase the peak inverse voltages to which rectifiers are subjected. The peak inverse voltage that a tube will withstand (especially gaseous tubes) varies with the operating temperature and a suitable safety factor should be used for variations of this type.

Inversion Circuits. An inverter is a device for converting direct current to alternating current. The oscillator discussed in Chapter X is such a device and is designed for high frequencies and generally for small amounts of power. For low frequencies and larger amounts of power, electronic devices such as the thyratron, the ignitron, and mercury-arc rectifiers are used in inverter circuits. The function of the tubes is to commutate or perform a switching operation. It is essential that the electronic device have grid control and that units of inductance or capacity or both be employed to produce inductive and capacitive storage capacity. Many combinations of these elements may be used for producing inversion.

One simple inverter circuit in which the device determines its own frequency is given in Fig. 19. In this circuit the resistance R_1 is relatively high and R_2 is relatively low. Before the d-c switch S is closed, the anode, grid, and cathode of the thyratron are at the same potential and the condenser C is discharged. When S is closed the grid is made positive with respect to the cathode, and hence if the d-c voltage is higher than the ionization potential, the thyratron will "fire" and current will flow through L, R_2, and C and R_1 in parallel. Since R_1 is large, the condenser will charge quickly and the potential of the

cathode will rise. The rise of potential of the cathode will tend to lower the potential across the cathode-anode circuit below the ionization potential. At the same time the inertia effect of the inductance L tends to maintain the current and will produce an induced voltage to maintain the current. As a result the condenser C will continue to charge and the voltage across it will rise above the d-c impressed voltage. When the energy stored in the core of L is expended, the voltage across the thyratron falls, leaving the potential of the cathode higher (owing to condenser) than the anode and higher than the grid.

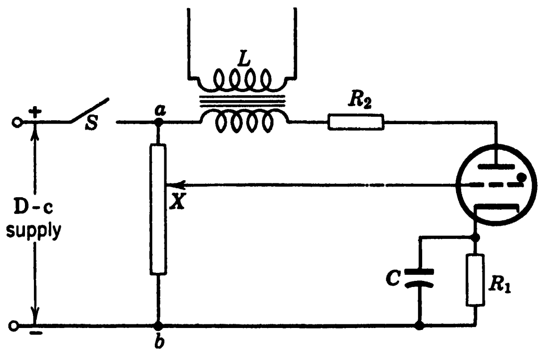


FIG. 19. Simple circuit using a thyratron for inversion.

Hence the current through the thyratron ceases and deionization takes place. In the meantime, the condenser discharges through the resistance R_1. As the condenser discharges and lowers its potential and that of the cathode to the point where anode-cathode potential reaches the ionization potential, and when the grid voltage reaches the point where it will fire the thyratron, the thyratron conducts current again, and the cycle is repeated. Obviously, the time of firing and the resulting frequency of the oscillating cycle can be varied by the position of X on the potentiometer ab which controls the grid bias. The frequency can be controlled also by variation of R_1 or C.

A transformer placed in the circuit of the thyratron would deliver alternating current to a suitable load. The wave form of this type of inverter is not sinusoidal.

In most power inverters the frequency is determined by a separate source. A simple schematic circuit for a single-phase inverter of this type is given in Fig. 20. Two thyratrons, A and B, with grids controlled by a separate voltage source are employed in a balanced circuit arrangement wherein both anodes are connected to the positive side of

the d-c supply. Assume that the grid supply at the instant of starting is positive for *A* and negative for *B*. Tube *A* fires and conducts electrons from the negative d-c line through the upper half of the transformer primary to the positive d-c terminal. The rise of current through the transformer winding creates an *IX* drop so that the potential of the anode of *A* falls to a value above zero equal to the cathode-anode drop in the tube. Under this condition the capacitor becomes charged, with the top plate negative and the lower plate positive. When the grid control potential reverses and makes grid of tube *A* negative and *B* positive, quick action follows. The negative grid

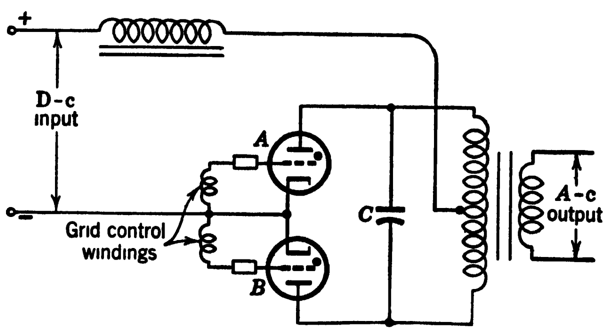


FIG. 20. Single-phase inverter using two thyratrons.

on *A* has no direct effect upon tube conduction, but the positive grid on *B* causes it to fire and the resulting surge of current through the lower half of the transformer winding lowers the potential of anode *B*. Simultaneously, electrons are supplied to the lower positive plate of the condenser *C*. This permits the unbound electrons on the top plate of *C* to flow off and lower the potential of the anode of *A*, thus permitting the now negative grid of *A* to regain control. At the end of the half-cycle the reverse process will take place with tube *A* conducting and tube *B* cut off. Thus the two tubes *A* and *B* serve as a reversing switch to cause alternate pulses of current to flow through the primary of the transformer. The resulting flux reversals in the iron core of the transformer gives an a-c output in the secondary having the same frequency as the applied grid potential.

Thyratrons have been used successfully for inversion by the General Electric Company on an experimental d-c transmission line between Mechanicville and Schenectady, New York. The inverter installation consisted of two six-phase units operating in a twelve-phase relationship. The inversion unit had a rating of 5250 kilowatts and

took constant current at 28,000 volts at full load from a d-c transmission line. This inverter installation was made in 1936 and has served to convert many millions of kilowatthours of electrical energy.

The multiphase mercury-arc rectifier also can be used as an inverter. A schematic diagram for this application is given in Fig. 21. This diagram is similar to that for rectification of alternating current

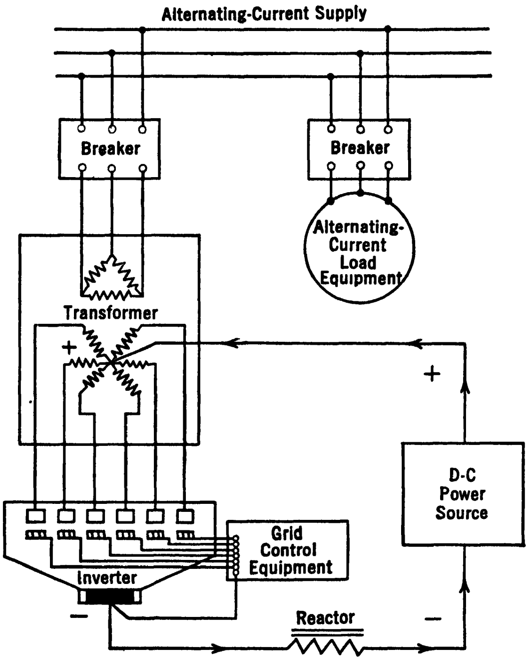


FIG. 21. Circuit using a mercury-arc rectifier for power inversion. (Courtesy Allis-Chalmers Manufacturing Company.)

but with three fundamental differences. First, the polarity at the d-c source is reversed though the direction of current flow through the rectifier is the same. Second, a reactor has been placed in the cathode lead. Third, the potential across the grid is controlled by a separate a-c source. The key to the inverter action lies in the application of and control of grid potentials. A positive potential on a grid will attract electrons to the grid and cause them to pass through to the anode. Hence if a-c potentials are applied to the multielectrode grids in the proper sequence, the resulting currents will pass through the secondary windings of the transformers connected to the anodes and produce a-c power in the primaries. The reactor in the cathode lead

smooths out the power flow from the d-c source and tends to reduce short-circuit current if an occasional short circuit should occur.

Ignitrons have been used for inversion and are destined to have an extensive application in this field. The circuit employed will follow that for the multiphase mercury-arc rectifier except that special firing circuits will be needed for the ignitrons. A twelve-phase ignitron rectifier which can be used for inversion is illustrated in Fig. 22.

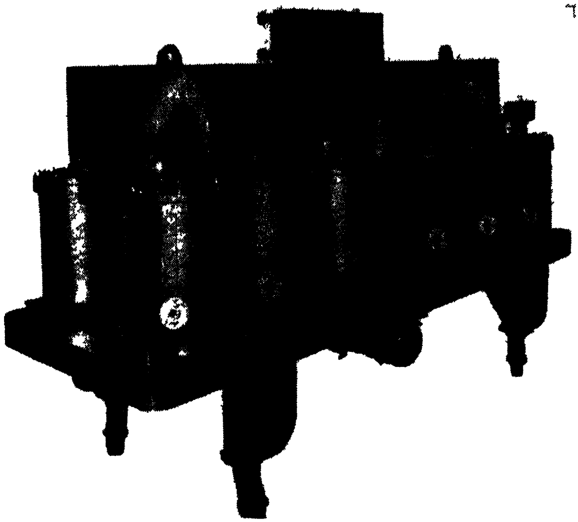


Fig. 22. Twelve-phase ignitron rectifier unit (Courtesy Westinghouse Electric Corporation.)

Possible applications of the mercury-arc rectifier as an inverter are for d-c transmission of power and conversion of d-c regenerated power on electrified railroads in mountainous districts.

Problems

1. A Tungar rectifier tube of Fig. 4, Chapter VI, is used for half-wave rectification as shown in Fig. 1 to charge a 6-volt storage battery for automobiles. If the battery takes 6 amperes at 7.5 volts, calculate the approximate a-c volts on the transformer secondary. If 115-volt alternating current is applied to the transformer, what is the approximate ratio of turns in the transformer? If the transformer has an efficiency of 80 per cent for this load, calculate the overall efficiency of rectification, taking care to include both cathode and arc losses.

2. Substitute a copper oxide rectifier for the tungar tube and recalculate

Problem 1, assuming a voltage of 12 volts effective per plate with one 1-volt drop in the forward direction.

3. Two phanotrons of Fig. 7, Chapter VI, are connected in the full-wave rectifier circuit of Fig. 2 to charge a 110-volt storage battery (125 volts when charging) at 40 amperes. Determine the approximate applied a-c volts on the secondary. If the transformer has an efficiency of 90 per cent, what is the overall efficiency of rectification?

4. The vacuum diode of Fig. 15, Chapter III, supplies a rectified current of 100 milliamperes to a resistance load of 2000 ohms. Calculate the ratio of the reduction of the voltage (equation 6).

5. The unfiltered current from a half-wave rectifier (sine wave applied) is passed through a d-c ammeter and a current transformer (1/1 ratio) connected to an a-c ammeter. When the d-c meter reads 12 amperes, what should be the indication on the a-c meter? Recalculate the problem for full-wave rectification?

6. A certain phase-shift voltage control on a three-phase rectifier delays the firing time to the peak of the impressed wave. What will be the resulting ripple factor? Consult Fig. 8 and neglect the effect of inductance in transformer and arc drop.

7. In the simple filter circuit of Fig. 10, assume that $C = 1$ microfarad, $R_L = 100{,}000$ ohms, and e_R falls for $\frac{1}{80}$ second. Determine from Fig. 33, Chapter X, the per cent of voltage drop from the maximum.

8. In an ideal half-wave rectifier circuit the impressed rms voltage is 65 volts. What is the d-c average voltage? If the peak value of current on the input side is 5 amperes, what is the I_{dc}?

9. In an ideal full-wave rectifier the rms values on the input are 90 volts and 6 amperes. Calculate the readings on a d-c voltmeter and ammeter on the load.

REFERENCES

Livingston and Lord, "The Single-Tube Thyratron Inverter," *Electronics*, April 1933.

Durand, S. R., *A Mercury-Arc Static Inverter*, Allis-Chalmers Mfg. Co., 1939.

M.I.T. Electrical Engineering Staff, *Applied Electronics*, John Wiley & Sons, 1943.

Chapter XIII

HIGH-FREQUENCY HEATING

Electric heating for industrial use may be produced by (1) resistance heating (I^2R), (2) induction heating, or (3) dielectric heating. Each of these methods differs in the theory involved and usually in the preferred applications. The first method is old in the art, the second has been utilized in some measure for half a century, and the third, dielectric heating, is a new application.

Most forms of resistance heating are independent of frequency and work equally well with direct current and at low voltages. Induction and dielectric heating require moderate or high values of frequency (also rather high voltages) and are classed as *high-frequency heating.* The higher frequencies require special methods of production in which electronic equipment is generally employed. Since resistance heating supplements the other forms of electric heating, a brief treatment of its theory and application will be given here.

Resistance heating is produced by the I^2R or power loss involved when an electric current flows through a heating resistor R. The heating resistor may be (1) a metal conductor, (2) a nonmetallic rod or tube of carbon, or (3) a liquid. Alloys of nickel and chromium are commonly produced in wire or ribbon forms to be wound in helical or coil forms for electric heating units. The heating units may be used in hot plates for heating glue or other compounds or in heat-insulated ovens for laboratory or production processes. Some ovens are furnaces for heat treating metals and other products. Sometimes the heating units are placed in ovens surrounding conveyer systems for drying and baking varnishes, paints, and enamels. For heat treating metals one type of furnace consists of carbon or Carborundum tubes (trade name Globar) through which high values of current pass to raise the temperature to red- or white-heat stage. Parts to be heated are placed directly within the heated tube. In other furnaces an insulated oven is heated by current flowing through solid Globar sticks.

Electrolytes and water (not distilled) can be heated by current

flowing between electrodes placed in the electrolyte. In this case the resistance of the liquid is the R for the I^2R heating. A similar application is employed in many chemical and metallurgical furnaces. Here the ores and other compounds that make the raw furnace charge are placed in the circular open furnace. Two or more large carbon or graphite cylindrical electrodes are lowered into the charge and current is passed from the electrodes through the charge. Initially, the charge is a solid which slowly fuses to a liquid mass under the action and control of electric currents of large magnitude. These furnaces are called carbon-arc furnaces.

A new and useful form of electric heat is produced by infrared (heat) bulbs built like the reflector type of incandescent lamp. These bulbs reflect and concentrate heat rays which may be directed on the work inside a conveyer type of oven. These heat lamps permit the concentration of heat on enamel or painted surfaces and obtain a quick drying action. The concentrated radiant heat penetrates the coating of enamel to a depth sufficient to produce rapid drying without wasting energy in heating the body of the work.

In passing it should be noted that the efficiency of conversion of electric energy into heat via resistance heating is 100 per cent. This means that all the electric energy delivered to the heating resistor is converted into heat energy. Not all this converted energy is delivered to the heating load but, where heat-insulated ovens surround the heating resistor, the ratio of heat delivered to the load to the converted heat energy may be high. Thus the overall efficiency in using resistance heating is likely to be high. This high overall efficiency permits resistance heating to be competitive with other methods of heating using fuels as a basis of heat energy.

Induction Heating. Heating by means of low-frequency magnetic fields has been known as far back as 1880. The first practical induction furnace was put into operation in the manufacture of electric-light bulbs and electron tubes for many years to fire "getters" and to remove imprisoned gas from metal and other parts within the tube. The application of high-frequency induction heating for the heat treatment of metals has been largely a wartime development.

The elementary theory of induction heating can be studied in Fig. 1. A coil of copper conductor called an inductor surrounds a metal cylinder called a charge or load. If an electric current is passed through the coil a magnetic field is created which passes through the cylinder. If the current is alternating, the field reverses direction for each cycle

and the flux linking the charge or cylinder changes accordingly. Such flux changes will produce heat losses within the charge. The nature and magnitude of these losses depends on the material and temperature of the charge. If the cylinder consists of a magnetic material such as iron, both hysteresis and eddy-current losses may occur. Hysteresis

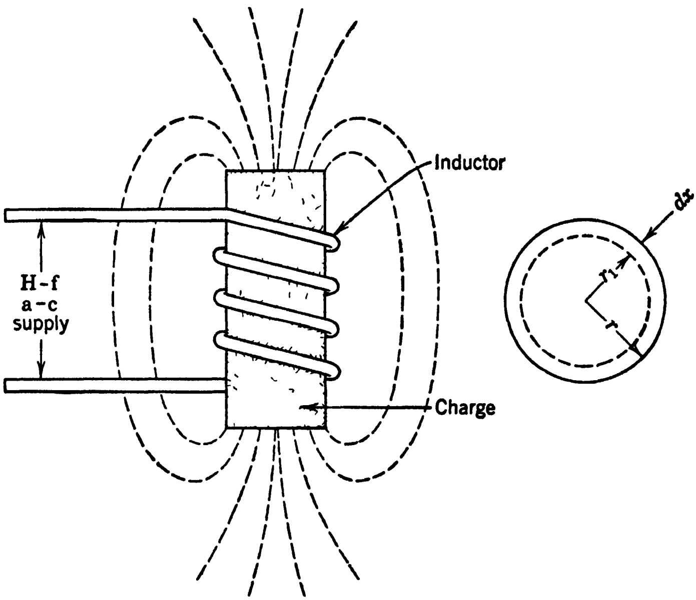


FIG. 1. Method of applying induction heating

losses are molecular-friction losses which follow the Steinmetz equation for hysteresis loss.

$$W = KB^{1.6}f$$

where W represents watts, B the flux density, and f the frequency. K is a constant to take care of the parameters involved.

It should be observed that the hysteresis loss varies directly as the frequency and as the 1.6 power of the flux density. Hysteresis losses are of major importance in the heating of iron and magnetic metals at temperatures below the Curie temperature (1420 degrees F for low-carbon steel). Above the Curie temperature magnetic properties effectively disappear and hysteresis loss ceases to exist. Thus it is easier to heat magnetic materials by the induction process but the limitation of hysteresis loss at higher temperatures and the fact that it does not apply to nonmagnetic metals frequently results in an omis-

sion of the effect of hysteresis in the development of some empirical equations for induction heating.

The theory of the production of eddy currents is suggested in Fig. 1. The view on the right is a cross section of the cylindrical charge. An elemental circular conducting ring or shell of thickness dx having an inside diameter r_1 and an outer diameter r may be considered. The changing flux through the charge links this elemental conduction shell and induces an emf in it where

$$e = -N\frac{d\phi}{dt}$$

This emf causes a circulating or eddy current to flow through the elemental shell like the current flow in the secondary of a transformer. The eddy current squared times the resistance of the path represents watts which are transformed into heat loss within the shell. An infinite number of elemental shell-like conductors may be visualized as taking part in the eddy-current heating.

Eddy-current losses can be expressed by the equation

$$W_e = Kf^2B_m{}^2$$

wherein the watt loss is proportional to the square of the *frequency* and the *maximum flux density*. Thus the total heat developed within the load can be controlled through the frequency and the flux density. The frequency employed has an important effect upon the depth of penetration of the heat because of a skin effect which takes place in the load. The distribution of the flux density, the current density, and the heat density varies greatly with the frequency employed, as illustrated in Fig. 2. This figure shows the approximate distribution for a charge consisting of a round stainless-steel bar for two different frequencies. In the upper part of the figure a low frequency is applied to the inductor coil (Fig. 1), so that a uniform flux density B results in the cylindrical charge. The emf induced along the axis of the cylinder by the changing flux will be zero. But, in circular shells surrounding the axis, emf's will be induced that are proportional to the enclosed flux which, in turn, is proportional to the square of the radius of the shell (area $= \pi r^2$). The eddy currents resulting from these emf's will depend on the resistance of the path ($I = E/R$). The resistance of the path is proportional to the length of path ($2\pi r$) and to r. Thus the current density at any point within the cylinder is proportional to emf/resistance or $r^2/r = r$. This gives the linear distribution for the

current density as shown in Fig. 2*b*. The heat produced by the eddy currents is proportional to the square of the current and hence to the square of the distance from the center *r*. This suggests the parabolic distribution of heat density shown in Fig. 2*c*. For a higher frequency of the power applied to the inductor, the eddy currents in the incremental shells of the charge will produce counter magnetomotive forces which will reduce the flux density near the center of the charge. This will result in a nonuniform distribution of flux density, as shown in

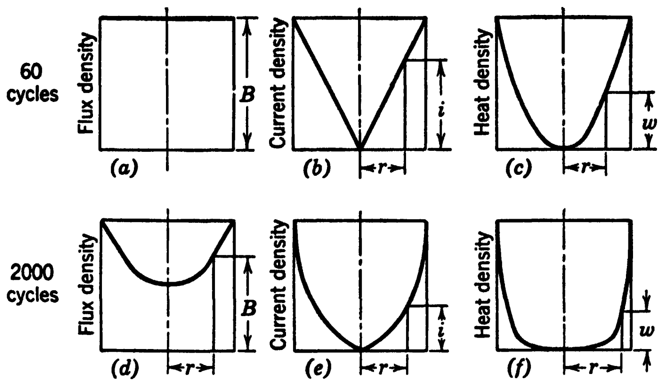


FIG. 2. (Courtesy Allis-Chalmers Manufacturing Company.)

Fig. 2*d*, and such nonuniform flux density will change the induced emf's and give a resulting current density of the form shown in Fig. 2*e*. The resulting heat-density distribution will be of the form shown in Fig. 2*f*. Thus the rise in frequency crowds the flux density and the heating to the surface of the charge or load. Obviously, the depth of heat penetration can be controlled by the applied frequency. The effectiveness of induction heating compared to flame or furnace heating is illustrated in Fig. 3.

Rigorous equations for eddy-current losses are quite complex and of little general use. Some approximate formulas do serve to show the relationship of the various parameters and hence are useful in analyzing induction heating. Unfortunately, these formulas cannot be applied to many practical problems because of the difficulty in measuring the magnitudes of the quantities involved. The following equation gives the power dissipated as heat in the surface of the charge.

$$\Delta P = \frac{\beta_t^2 \sqrt{\rho \mu f}}{8\pi} \tag{1}$$

where ΔP = power dissipated by eddy currents per unit of volume.
β_t = tangential component of magnetic flux density at surface of charge.
ρ = resistivity of the charge.
μ = permeability of charge (unity if nonmagnetic).
f = frequency.

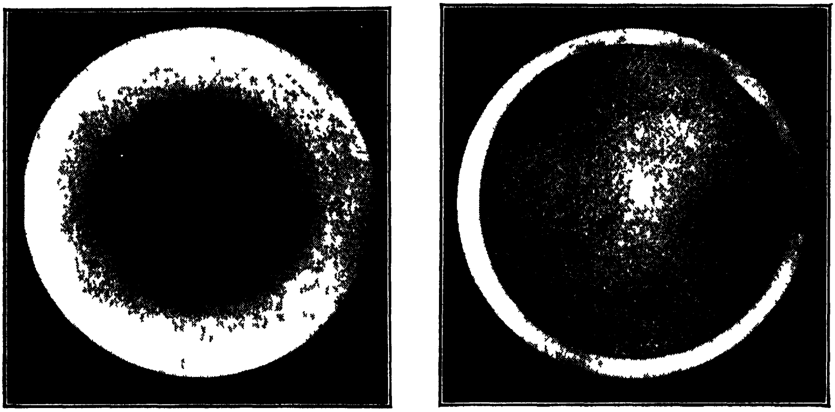


FIG. 3. Comparison of the heating of a shaft with a forge (*left*), and the induction processes (*right*) (Courtesy Allis-Chalmers Manufacturing Company)

Since the magnetic flux density (β_t) is proportional to the ampere-turns in the coil (for nonmagnetic materials), the factor β_t^2 can be replaced by I^2N^2 (times a constant).

The depth of penetration of the heat may be expressed by the following equation where depth δ refers to the point at which the eddy currents fall to a value equal to $1/e$ (37 per cent) times their magnitude at the surface.

$$\delta = \frac{1}{2\pi}\sqrt{\frac{\rho}{\mu f}} = \text{depth of penetration} \tag{2}$$

Combining equations (1) and (2) gives

$$\Delta P = \frac{\beta_t^2 \rho}{16\pi^2 \delta} \tag{3}$$

The following analysis of equations (1), (2), and (3) summarizes some of the salient points in the application of induction heating.

1. For a given flux density, frequency, and resistivity, the heat input per unit volume is proportional to the square root of the per-

meability (equation 1). Thus magnetic materials will heat more readily than those that are nonmagnetic and have a permeability of one.

2. For a given material, the depth of heat penetration is inversely proportional to the square root of the frequency (equation 2). Thus the depth of penetration may be controlled by the chosen frequency.

3. For a given flux density and depth of penetration, the heat input

FIG 4 Parts selectively hardened by electronic heaters. (Courtesy General Electric Company)

per unit of volume varies as the resistivity (equation 3). Thus materials with high resistivity can be heated more readily.

4. For a given material and frequency, the heat input per volume is proportional to the square of the flux density (equation 1). Thus the heat input can be controlled by varying the number of turns in the inductor or the current through the inductor or both

Induction heating is used for melting, annealing, hardening, brazing, and soldering of metals. Some induction furnaces for melting metals have used frequencies as low as 25 to 60 cycles. For heating forgings, for annealing, and for deep surface hardening where larger pieces are involved, frequencies in the range of 500 to 3000 cycles generally give the highest efficiencies. For small pieces, for very thin

core hardening of steel, for brazing and for soldering, high frequencies in the range of 100,000 to 1,000,000 cycles are used. Some examples of parts that have been selectively hardened are shown in Fig. 4 and others that are being inductively brazed in Fig. 5.

Some of the methods by which induction heating may be applied to machine parts are shown in Fig. 6. Part *a* shows the method of applying surface heating to a shaft. If the coil covers only a short

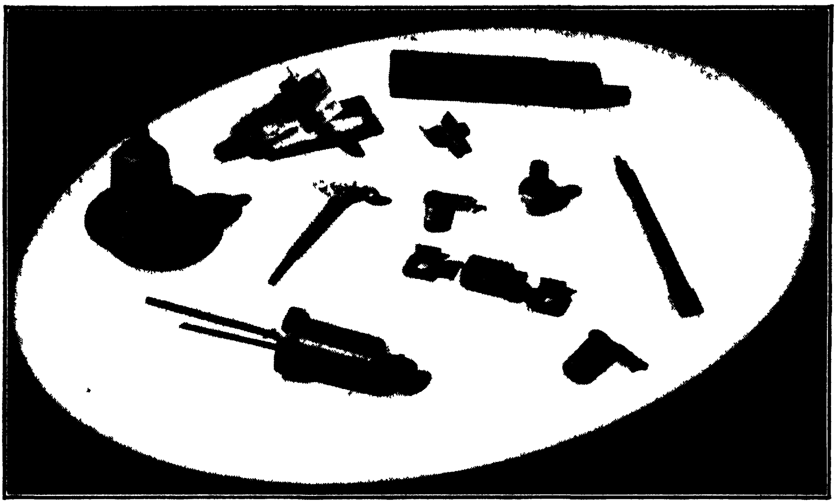


FIG. 5. Parts inductively brazed with a 5-kilowatt electronic heater. (Courtesy General Electric Company.)

portion of the shaft, that portion may be surface heated and hardened without affecting the interior or surface of the rest of the shaft. The direction of current flow in the charge (cylinder) for an increasing current in the inductor is shown in Fig. 6*a*. If the charge is moved outside of the inductor as shown in *b* and *c*, a current will flow in the charge in the same direction as before, but the magnitude of the current will be smaller because the flux density is much lower outside the coil than within. Parts *d*, *e*, and *f* of Fig. 6 and the jig of Fig. 7 illustrate the application of induction heating to actual heating problems. Several methods of applying induction heating for brazing and soldering operations are illustrated in Fig. 8.

A radio-frequency coil surrounding a thin metal plate will induce heat into the surface of that plate. This principle was applied during the war period to melt a very thin layer of tin on tin plate sheets and

thus cause it to flow, giving a very thin and complete coverage of the surface. The development of this process in the research laboratories of the Westinghouse Electric Corporation helped to conserve the limited stock pile of tin reserves in the United States.

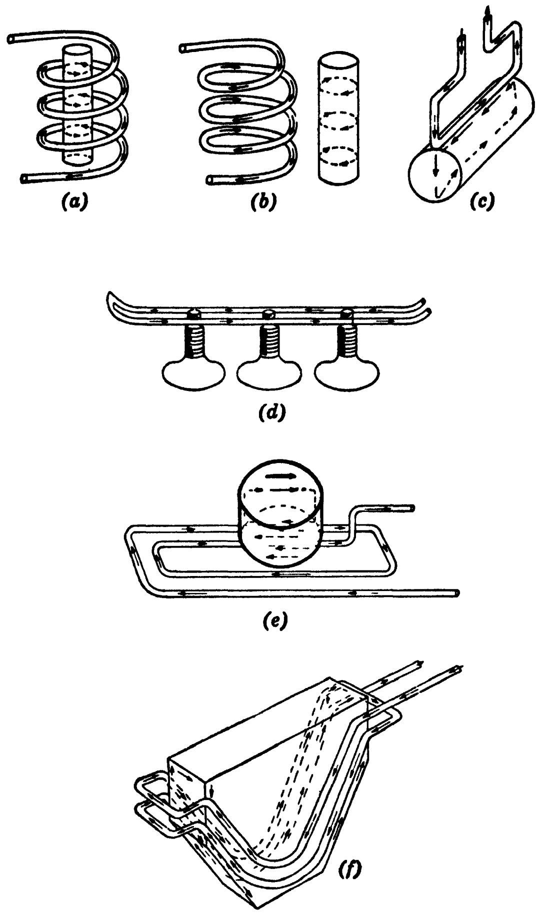


FIG. 6. Sketches showing the instantaneous direction of current flow in the part adjacent to the current in an inductor coil as used in electronic heating. (Courtesy General Electric Company.)

Some of the advantages of induction heating which have brought about its wide adoption in the metal-working industry are as follows:

1. Speed of heating metals with consequent reduction of oxidation and scaling.
2. Ability to heat limited parts or surfaces of a metal piece instead of an entire piece.

3. Use of power for heating only when desired, without necessity of maintaining furnace temperature continuously.
4. Thorough mixing of alloys when melting due to the "stirring effects."
5. No contamination of the charge by fuel gases.
6. Better working conditions for the operators than with fuel- or gas-fired furnaces.

The overall efficiency of high-frequency induction heating is the ratio between the heat generated in the charge or load and the energy taken from the supply line. This efficiency is rather low because two factors or efficiencies are involved, neither of which is likely to be high. First, there is the efficiency of conversion from a frequency of 60 cycles to that required for the heating process. Second, there is the inductor efficiency or the ratio of the heat generated in the charge to the input to the inductor. Considering for the moment the inductor efficiency only, it may depend on the frequency and the size or the diameter of the charge. It may be inferred from equation 1 that the efficiency would rise with the frequency. This is generally true for small parts and for small depth of heat penetration, but it is not true for charges of larger diameter. The effect of an increase in the diameter of charge can be analyzed as follows. For a given frequency and inductor current, let the radius of load be increased. The area of load, the induced flux, and the induced emf's will vary as the square of the radius. The average resistance of the eddy-current path will remain constant because the length and cross section of the path will increase in the same ratio. Since the heat generated is a function of E^2/R, the generated heat will tend to vary as the fourth power of the radius of the charge. However, as the radius of the charge increases, the eddy current will reduce the flux density inside the charge so that the emf's will not rise as the fourth power and the heat loss in charge will tend to become linear. This relationship is illustrated in Fig. 9 where the dotted line is plotted for the fourth power. The actual heat generated (line Y) starts as the fourth power and changes to a linear relationship. Line X gives the heat loss in the inductor which is linear since inductor

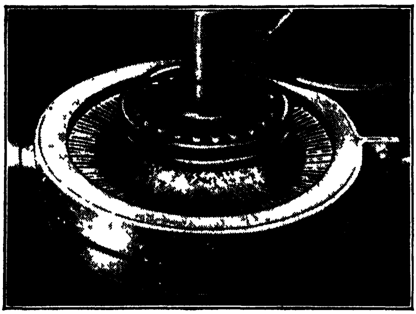


FIG. 7. Induction heating and water quenching for hardening the face of gear teeth. (Courtesy Westinghouse Electric Corporation.)

length varies as r. It is obvious from Fig. 9 that the inductor efficiency is very low for small diameters of charge but rises to a constant value after the heat in the charge becomes linear. This relationship is

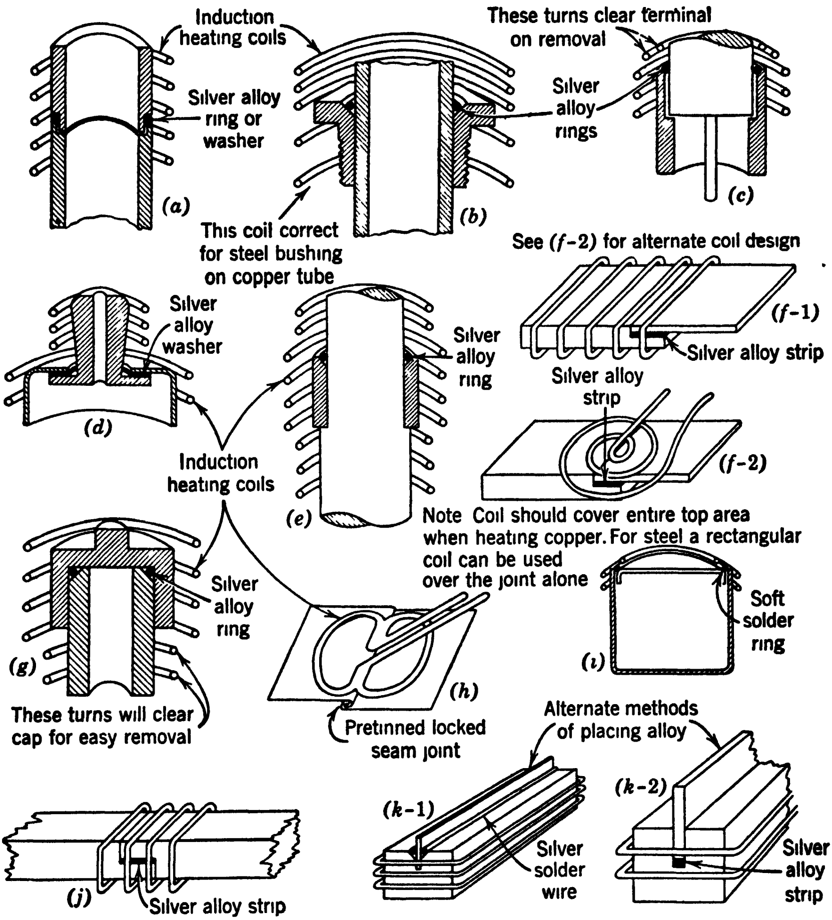


FIG. 8. Typical joints and heating coil locations sketched for brazing and soldering by induction heating. (Courtesy General Electric Company.)

also shown in Fig. 10 which applies to an aluminum cylinder in which the diameter is equal to the height. Two other important considerations are illustrated in Fig. 10. First, the maximum efficiency attained for large diameters of load is the same regardless of the frequency used (65 per cent for the case illustrated). Second, the optimum frequency for a given job depends on the diameter of the charge. For large

diameter of charge the lower frequencies of 1000 to 1500 cycles give the same inductor efficiency as a higher efficiency with the advantage that the lower frequency can be produced at a lower cost and with a higher efficiency of conversion.

In the design, construction, and use of "inductor" coils for induction heating, the following suggestions will be useful.

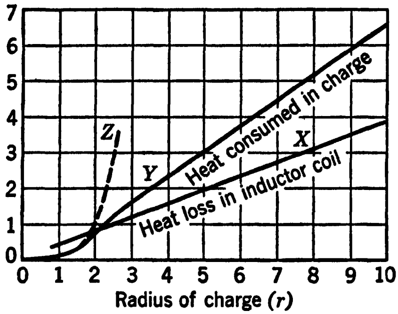


FIG. 9. (Courtesy Allis-Chalmers Manufacturing Company.)

1. The heat generated in a part is due entirely to the magnetic flux created by the inductor coil. The flux density is greatest at the conductor itself; hence the closer the coil is to the part being heated, the greater will be the heating.

2. The coil should conform to the shape of the part if no sharp corners need be considered.

3. Sharp contours will heat first because of the concentration of flux and lack of mass. Thus the coil should be the farthest from the part at these points.

4. In heating dissimilar metals for brazing, the magnetic flux must be concentrated on the slowest heating metal and the joint should be at the correct temperature before the brazing alloy melts so that it will be drawn into the joint. Thus a concentration of heat on the brazing alloy should be avoided.

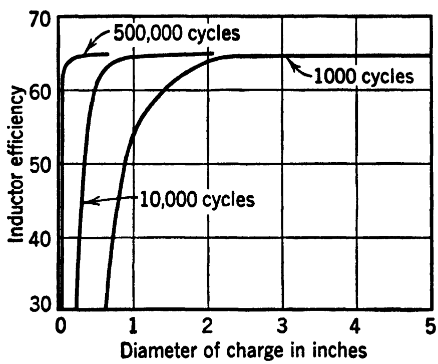


FIG. 10. (Courtesy Allis-Chalmers Manufacturing Company.)

5. In hardening, double-bank (double-layer) coils are sometimes necessary because of the current limitations of electronic heaters. (The outside layer of turns is far less efficient and such coils should be used only when necessary.)

6. Uniform heating at the circumference of a disk or bar may be secured by rotating the parts to avoid the heating effects of coil leads.

7. A highly concentrated band of heat on a part or for zones that

are not readily accessible may be secured by the use of a single-turn coil carrying a high current. Since it is not economically feasible to construct electronic heaters with a circulating current above 300 amperes, it is necessary to use output transformers. Such transformers, being of the air-core type, have a poor coupling and a low overall efficiency.

Calculations for induction heating to determine the thermal power, input power, and inductor design are very difficult to perform. Equations following the basic theory may not be usable because of the many varying factors entering into the application. Here, as is so often true in electrical design, empirical formulas and curves deter-

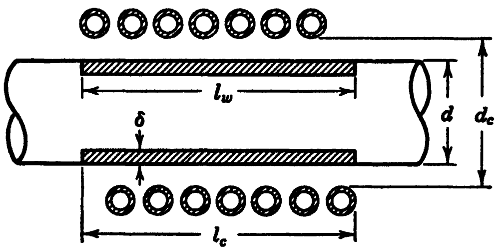


FIG. 11

mined from laboratory experiments and practice must be employed. Hence the industrial user of induction-heating equipment relies upon the recommendations of the design specialists of the equipment manufacturer. To introduce the reader to the complexities of induction-heating calculations, the following example is drawn with permission from the book *Electronics for Industry* by W. I. Bendz and C. A. Scarlott, John Wiley & Sons, 1947.

A solid cylindrical bar of metal is to be heated by an inductor coil as shown in Fig. 11. The cylinder will be heated over a length of l_w and, as a result of skin effect, the metal will be heated for a short depth on the surface as suggested in Fig. 2, part *f*. For the purpose of calculation it may be assumed that all the heat is confined to a fictitious layer (shaded area of Fig. 11) having a thickness δ called the *depth of penetration.* The depth of penetration may be calculated by the empirical equation

$$\delta = 3170 \sqrt{\frac{\rho}{\mu f}} \text{ inches} \tag{4}$$

where ρ = resistivity (ohm-inches).
μ = magnetic permeability.

The induced current flows through a path whose resistance is

$$\text{Resistance of path} = \frac{\rho P}{\delta l_c} = R_w$$

where P = average length of path (inches). The current through the work I_w acts as a single-turn secondary of a transformer; hence. neglecting losses,

$$I_w = \text{coil amperes} \times \text{coil turns} = I_c T_c$$

where I_c = inductor current.
T_c = inductor turns.

The total power input to the load is

$$\text{Total power input} = I_w{}^2 R_w = (I_c T_c)^2 R_w = (I_c T_c)^2 \frac{\rho P}{\delta l_c} \tag{5}$$

The power input can be reduced to watts per square inch by dividing by the heating surface area Pl_c and, if turns of the inductor is changed to turns per inch N, equation 5 reduces to

$$\text{Watts per square inch} = (I_c N)^2 \frac{\rho}{\delta}$$

Substituting the value of depth of penetration from equation 4 gives

$$\text{Watts per square inch} = 3.16 \times 10^{-4} (I_c N)^2 \sqrt{\mu \rho f} \tag{6}$$

For all nonferrous metals, such as aluminum, brass, and copper, the permeability is approximately unity so that a substitution of $\mu = 1$ can be made in the preceding equations. Empirical data for iron at normal saturation and at temperatures below the Curie point gives

$$\mu = 1.8 \frac{B}{H}$$

and for a normal saturation of B = 18,000

$$\mu = \frac{32{,}400}{H} \tag{7}$$

For iron and steel at temperatures above the Curie point $\mu = 1$.

The power required to raise the temperature to the desired final temperature within a specified heating time is known as the *thermal power*. It is an expression for the rate of heating and does not include losses of any kind. Thermal power is calculated from the formula

$$P_T = 17.6Mc\,\Delta T \text{ watts} \tag{8}$$

where M = rate of material heated in pounds per minute.
c = specific heat of material.
ΔT = temperature rise in degrees Fahrenheit.
P_T = thermal power in watts.

The work or charge heated by induction is subject to heat losses by radiation, convection, and conduction. These losses must be supplied by the input to the inductor coil. These losses may be calculated by empirical formulas or determined in the laboratory and given in curves like those of Fig. 12.

Power density is the total power input to the material heated divided by the volume of metal within the coil. The total power input is the sum of the thermal power from equation 8 plus the radiation and conduction losses.

$$PD = \frac{\text{Total power input (watts)}}{\text{Volume of metal in coil (cubic inches)}} \text{ watts per cubic inch} \tag{9}$$

The calculation of the peak magnetizing force H required at the surface of a work piece is a hazardous factor in induction-heating problems. The following are among the many formulas that have been used.

Solid magnetic cylinder

$$H = \left(\frac{PD \times d}{0.438\sqrt{f\rho}}\right)^{2/3} \tag{10}$$

Nonmagnetic cylinder

$$H = \sqrt{\frac{PD \times d}{2.59 \times 10^{-3}\sqrt{f\rho}}} \tag{11}$$

where d is the diameter of the cylinder. Equations 10 and 11 are valid only when

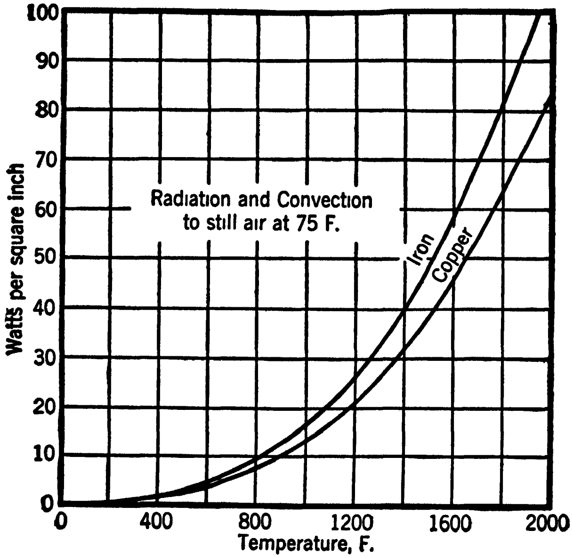
$$d\sqrt{\frac{\mu f}{\rho}} \geqq 13{,}400 \tag{12}$$

The value of ρ used in the equations should correspond to the average temperature of the material being heated. This average value can be obtained by

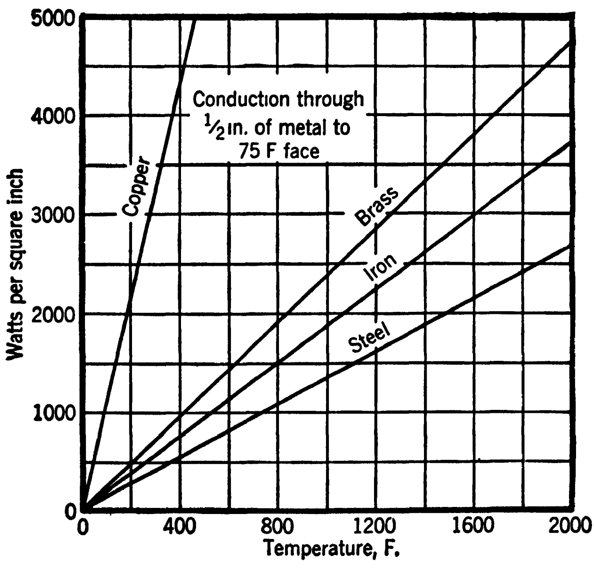
$$\rho_{\text{hot}} = \rho_{70^\circ\text{F}}\,[1 + \alpha(T_{\text{hot}} - 70)] \tag{13}$$

The coil current for multiturns in the inductor is computed by the following relation:

$$I_c = \frac{1.43 \times H}{N} K_1 \tag{14}$$



(a)



(b)

FIG. 12. Typical radiation and convection loss (a) and conduction (b) through ½ inch of metal. (Reproduced with permission from W. I. Bendz and C. A. Scarlott, *Electronics for Industry*, John Wiley & Sons, 1947.)

or

$$I_c = \frac{1.43 \times H \times l_c}{T_c} K_1 \tag{15}$$

where K_1 is an empirical factor given in Fig. 13 which depends upon the ratio of coil length to the spacing between the coil internal diameter and the piece to be heated.

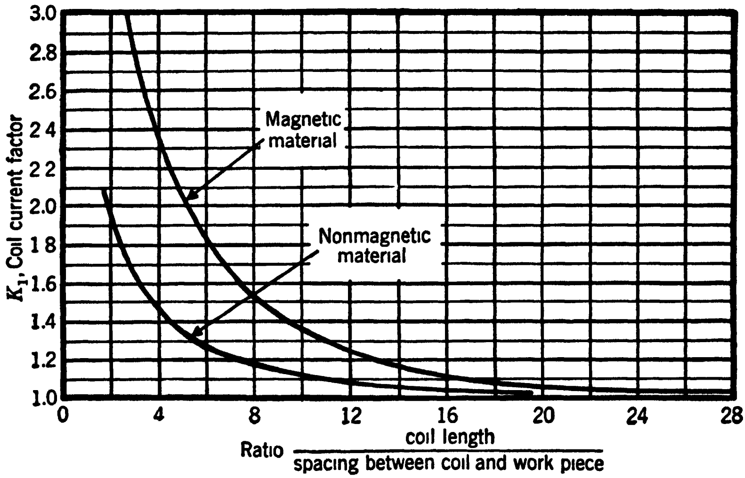


FIG. 13. The coil-current K_1 factor is a function of the ratio of the coil length and the spacing between the coil internal diameter and the work piece. (Reproduced with permission from W I Bendz and C A Scarlott, *Electronics for Industry*, John Wiley & Sons, 1947)

The voltage induced in the inductor by the flux within the coil is

$$E_i = \sqrt{2}\pi f T_c \phi_0 \times 10^{-8} \text{ volt (rms)} \tag{16}$$

The total flux ϕ_0 is composed of the flux in the air space between the coil and the work and the flux in the work. In addition to the induced voltage from the flux, the *RI* voltage drop within the coil itself must be added to obtain the total coil voltage. The voltage induced in the inductor because of the flux in the air space is

$$E_a = 28.6 f T_c H(A_c - A_w) \times 10^{-8} \text{ volt} \tag{17}$$

where A_c = the inside area of the coil (square inches).
A_w = the area of the work piece (square inches).

The voltage caused by flux in the work of nonmagnetic material may be obtained approximately from

$$E_w = 28.6fT_cH[A_w(P - jQ)]10^{-8} \text{ volt} \tag{18}$$

where P and Q have the following values:

Nonmagnetic cylinder

$$P = Q = \frac{6310}{d\sqrt{f/\rho}} \tag{19}$$

where ρ = resistivity of material at maximum temperature and when $d\sqrt{f/\rho} \geqq 13{,}400$.

When the work piece consists of magnetic material at a temperature below the Curie point, it may be assumed that the flux penetrates to a depth δ_f below the surface and that the flux density is $B = 18{,}000$. On this basis the total flux in the work piece is $18{,}000\pi d\delta_f$, where

$$\delta_f = 17.6\sqrt{\frac{H\rho}{f}} \tag{20}$$

and ρ = resistivity of material at maximum temperature. Now the voltage induced by the work is

$$E_w = 0.0162dfT_c\delta_f \text{ volts} \tag{21}$$

The voltage drop across the inductor is the product of the coil current and the effective a-c coil resistance R_c for frequency used. The calculation of R_c becomes somewhat involved because of the leakage flux which links the turns. The voltage drop may be assumed to be

$$E_{ir} = I_cR_c + jI_cR_c \tag{22}$$

$$\text{Coil copper loss} = I_c^2R_c \tag{23}$$

Summarizing the preceding discussion it is evident that the total power load on the radio-frequency generator will be:

1. The thermal power input to the work (equation 8).
2. Radiation, convection, and conduction losses (Fig. 12).
3. Coil copper loss (equation 23).

The following example will illustrate the application of the preceding empirical formulas, curves, and theory to an induction-heating problem.

Assume that the ends of steel bars 1 inch in diameter are to be heated for 6 inch. The initial temperature is 70 degrees F and the bars are to be heated to 1200 degrees F. To satisfy the production rate required, each piece must be heated in 30 seconds. A frequency of 450 kilocycles is to be used.

The physical properties of the metal, selected from Table 1, are as follows:

Density—0.285 pound per cubic inch.
Specific heat (average)—0.118 Btu per pound per degree Fahrenheit.
Resistivity—6.75×10^{-6} ohm-inch at 70 degrees F.
Temperature coefficient of resistance—0.0024 per degree Fahrenheit.

1. Thermal Power.

$$\text{Volume of metal enclosed by coil} = \frac{6\pi(1)^2}{4} = 4.71 \text{ cubic inches}$$

$$\text{Rate of heating} = \frac{4.71 \times 0.285}{0.5} = 2.68 \text{ pounds per minute}$$

$$\begin{aligned} P &= (17.6)(2.68)(0.118)(1200 - 70) \\ &= 6310 \text{ watts} \end{aligned} \tag{8}$$

2. Radiation, Convection, and Conduction Loss.

Radiation and convection, from Fig. 12*a*	= 5	watts per square inch
Conduction loss, from Fig. 12*b*	= 800	watts per square inch
Total radiation surface = $\pi(1)(6)$	= 18.9	square inches
Conduction surface = $0.7854(1)^2$	= 0.79	square inch
Radiation and convection loss = 5(18.9)	= 95	watts
Conduction loss = 0.79(800)	= 632	watts
Total	727	watts

3. Power Density.

$$PD = \frac{\text{Watts input to work piece}}{\text{Volume of metal in coil}} = \frac{(6310 + 727)}{4.71}$$

$$= 1495 \text{ watts per cubic inch} \tag{9}$$

4. Peak Magnetizing Force.

The peak magnetizing force is found by applying equation 10 to a solid magnetic cylinder:

$$H = \left(\frac{PD \times d}{0.438\sqrt{f\rho}}\right)^{2/3}$$

The value of ρ corresponding to the average temperature, 600 degrees F, is first determined by

$$\begin{aligned} \rho_{600} &= 6.75 \times 10^{-6}\ 1 + 0.0024(530) \\ &= 15.3 \times 10^{-6} \text{ ohm-inch} \end{aligned}$$

$$\begin{aligned} H &= \left(\frac{1495 \times 1.0}{0.438\sqrt{450{,}000 \times 15.3 \times 10^{-6}}}\right)^{2/3} \\ &= 118 \text{ oersteds} \end{aligned}$$

The validity of equation 10 is then checked by equation 12:

$$d\sqrt{\frac{\mu f}{\rho}} \geqq 13{,}400$$

where

$$\mu = \frac{32{,}400}{H} = \frac{32{,}400}{118} = 275 \tag{7}$$

Then

$$1.0\sqrt{\frac{(275)(450{,}000)}{15.3 \times 10^{-6}}} = 9.0 \times 10^{6}$$

This is greater than 13,400, and equation 10 is valid.

5. Assumed Coil Design.

Assume the coil to be wound of ¼-inch outside-diameter copper tubing, three turns per inch, 6.0 inches long, 18 turns, 1.75-inch pitch diameter or 1.50-inch inside diameter.

$$\text{Spacing between coil inside diameter and work piece} = \frac{1.5 - 1.0}{2}$$

$$= 0.25 \text{ inch}$$

$$\text{Ratio } \frac{\text{length}}{\text{spacing}} = \frac{6.0}{0.25} = 24$$

6. Coil Current.

From Fig. 13, if we use ratio $\dfrac{\text{length}}{\text{spacing}} = 24$, the coil current factor K_1 is 1.05. From equation 14 the coil current is

$$I_c = \frac{1.43(118)}{3.0} \times 1.05 = 59.0 \text{ amperes} \tag{14}$$

7. Coil Voltage.

(a) Voltage due to Flux in the Air Space.

Coil cross-sectional area	$= 0.7854(1.5)^2$	$= 1.76$ square inches
Work piece area	$= 0.7854(1.0)^2$	$= 0.79$ square inch
Area of air space		$= 0.97$ square inch

$$E_a = 28.6(450{,}000)(18)(118)(0.97) \times 10^{-8}$$

$$= 266 \text{ volts} \tag{17}$$

(b) Voltage due to Flux in Work Piece.

$$\delta_f = 17.6\sqrt{\frac{H\rho}{f}} \tag{20}$$

Where resistivity ρ is at maximum temperature

$$\rho_{1200} = 6.75 \times 10^{-6}\,[1 + 0.0024(1200 - 70)]$$
$$= 25.1 \times 10^{-6} \text{ ohm-inch}$$

$$\delta_f = 17.6\sqrt{\frac{(118)(25.1 \times 10^{-6})}{0.450 \times 10^6}} = 0.00143 \text{ inch}$$

$$E_w = 0.0162(1.0)(450{,}000)(18)(0.00143)$$
$$= 188 \text{ volts} \tag{21}$$

(c) Voltage due to Coil Resistance and Flux in Turns.

Length of tubing in coil $= \pi 1.75 \times 18 =$	99.0 inches
Length of tubing in leads to coil $=$	8.0 inches
Total	107.0 inches

A-c resistance of tubing $= 785 \times 10^{-6}$ ohm per inch
A-c resistance of coil $= (785)(107)10^{-6} = 0.084$ ohm

$$E_{ir} = (59)(0.084) + j(59)(0.084)$$
$$= 4.95 + j4.95$$
$$= \sqrt{(4.95)^2 + (4.95)^2}$$
$$= 7.0 \text{ volts} \tag{22}$$

(d) Total Coil Voltage.

1. Due to flux in air space	266 volts
2. Due to flux in work piece	188 volts
3. Due to coil resistance	7 volts
Total, E_c	461 volts

8. Coil Copper Loss.

$$\text{Copper loss} = (59)^2(0.084)$$
$$= 292 \text{ watts} \tag{23}$$

9. Total Power Load.

Thermal power input to work piece	6310 watts
Radiation, convection, and conduction loss	727 watts
Coil copper loss	292 watts
	7329 watts

10. Total Volt-Amperes.

$$\text{Load volt-amperes} = (461)(59) = 27{,}200 \text{ volt-amperes}$$

$$\text{Ratio } \frac{\text{load volt-amperes}}{\text{load watts}} = \frac{27{,}200}{7{,}329} = 3.72$$

11. Generator Rating.

Required power output	7 3 kilowatts
Required radio-frequency current	59 amperes
Required volt-amperes	27 2 kilovolt-amperes

These values are all conservatively available by selecting a 10-kilowatt 450-kilocycle radio-frequency generator.

TABLE 1 PHYSICAL CONSTANTS OF METALS

Material	Resistivity (ohm-inches at 70° F)	Temperature Coefficient of Resistance (degrees F)	Average Specific Heat (Btu per pound per degree F)			Thermal Conductivity (Btu per hour per square foot per inch per degree F)	Weight (pounds per cubic inch)
			70° F	200° F	1200° F		
Aluminum	1 10 $\times 10^{-6}$	0 00220	0 214	0 230	0 274	1392	0 098
Brass, 60–40 (Muntz metal)	2 39 $\times 10^{-6}$	0 00109	0 092	0 094	0 100	877	0 303
Copper	0 679 $\times 10^{-6}$	0 00218	0 092	0 094	0 110	2668	0 321
Iron, pure	3 86 $\times 10^{-6}$	0 00361	0 104	0 113	0 210	467	0 282
Nickel	3 07 $\times 10^{-6}$	0 00334	0 105	0 113	0 126	415	0 322
Steel							
Soft	6 75 $\times 10^{-6}$	0 0024	0 110	0 114	0 143	334	0 284
Tempered hard	18 0 $\times 10^{-6}$	0 00089	0 120	0 125	0 150	334	0 283

Dielectric Heating. Dielectric heating is the term applied to the generation of heat in nonconducting materials by their losses when subjected to an *alternating* electric field. This form of heating is applied by placing the nonconductor load between two electrodes across which a high-frequency voltage is impressed as shown in Fig. 14. The

load and electrodes constitute a capacitor in effect and a current flows in the circuit in accordance with the equation $I_c = E/X_c$. In an *ideal* capacitor the current leads the impressed voltage by 90 degrees, zero power is absorbed in the capacitor, and the power factor (pf) is zero. It is impossible to construct an ideal capacitor and those capacitors which are produced for industrial use do have some losses and some heat is generated. In the dielectric heating or load capacitor (Fig. 14), the losses are much higher and are utilized as the source of heat. The cause of the internal heating is not clearly understood but appears to arise from the following sources. First, a small leakage current flows through the material due to the a-c voltage impressed across it. This current produces an I^2R heat loss. Second, the alternating potential gradient in the nonconducting material results in a dielectric hysteresis which is a form of molecular friction analogous to the magnetic hysteresis in magnetic materials. Since heating is the desired effect, differentiation as to the cause is not important. The power factor of the material used for a load is the ratio of the heat generated to the volt-amperes input to the load.

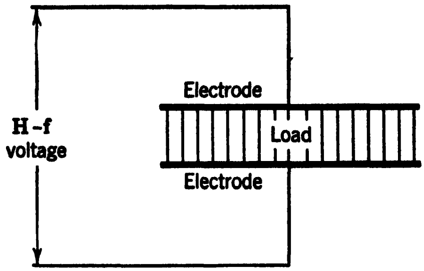


FIG. 14. Method of applying dielectric heating.

$$\underset{\text{Power factor}}{(pf)} = \frac{\overset{(W)}{\text{Heat generated}}}{\underset{(EI)}{\text{Volt-amperes}}}$$

and

$$W = EI \times pf$$

But

$$I = \frac{E}{X_c} = 2\pi fCE10^{-6} \text{ (approximately)}$$

and

$$W = 2\pi fCE^2 10^{-6} \times pf \qquad (24)$$

where C is in microfarads.

Equation 24 shows that, if all other factors are held constant, the heat generated in the load will vary directly as (1) the frequency, (2)

the capacity of the heating load, (3) the voltage squared, and (4) the power factor of the load. Obviously, the frequency and the voltage are directly controllable by the operator of the heating process. The capacity of the load will be determined by the geometry of the electrodes and the applied load and by the dielectric constant of the material used in loading. The power factor of the load will depend on the material of the load.

In contrast to induction heating, the heat loss in the insulating or nonconducting material used for dielectric heating may be calculated with a fair degree of accuracy. The calculation for the heat generated follows directly from equation 24. The capacity of the load may be calculated from the following equation:

$$C = \frac{2248AK}{10^{10}d} \text{ (for parallel plates)} \tag{25}$$

where C = capacity in microfarads.
A = area of one electrode in square inches.
d = distance between electrodes in inches.
K = dielectric constant.

The heat energy required for a given heating load may be determined from equation 8. In applying the above equations to practical problems the following precautions should be observed.

1. The power factor of most materials varies with frequency and temperature. Thus, the power factor should be measured near the operating frequency and at a known temperature.

2. The dielectric constant changes relatively slowly with frequency and temperature and can be taken from published tables with fair accuracy.

3. The equations hold only for uniform electric fields. Since the field at the edges of the electrodes is always distorted, a good approximation is possible only when the minimum dimension of the electrode is large compared to the distance between the plates. The use of plates somewhat larger than the material to be heated will reduce nonuniform heating at the edges.

4. To avoid corona and arcing effects, the maximum voltage applied to the electrodes is approximately 14 to 15 kilovolts rms, with 2 to 3 kilovolts rms per inch of separation as a maximum for smaller spacings.

5. The maximum electrode dimensions must be limited to at least one-eighth of the wavelength to avoid standing waves which result in nonuniform heating. The formula for wavelengths is

$$\text{Wavelength (meters)} = \frac{300}{f\ \text{(megacycles)}}\ \text{(in air)} \qquad (26)$$

6. The charge to be heated must be of uniform analysis throughout and must contact each plate. An air gap between the electrodes and the charge results in a series-capacitor effect and introduces serious errors in the above equations. (It is sometimes desirable to employ electrodes designed to provide an air gap between electrodes and the charge. In such cases the voltages employed must be much higher to provide the same potential gradient across the charge as would exist with contact electrodes.)

The frequency employed for dielectric heating depends on the size of the charge and the power required. For some materials a very high frequency may be desired. The equipment available limits the frequency to approximately 200 megacycles for power outputs of 100 watts. At 30 megacycles power up to 40 kilowatts is available and for 4 megacycles much higher power ratings are available.

The dielectric constants for most materials fall in the range of 2 to 6, but they may vary from 1 for gases up to 1000 for some ceramics. The power factors found in dielectric heating usually lie between 0.02 and 0.07, but they may be as low as 0.00015 (mica, polystyrene) or as high as 0.15 (asbestos). Gases and pure water have power factors that are essentially zero and cannot be heated.

The important advantage of dielectric heating lies in the fact that the heat is generated inside the material. With a uniform material and a uniform alternating electric field, the heat generation is the same and the rise of temperature is uniform throughout the charge. This uniformity of temperature rise also gives *speed* of heating. Any conventional method of heating will require the application of heat from the sides of the charge. Since nonconducting materials usually have poor heat conductivity or transfer, such heating is necessarily slow and results in nonuniform temperatures.

There have been an increasing number of applications of dielectric heating to nonmetals and nonconducting materials. Some of these applications are:

1. Gluing, drying, and curing of wood.
2. Preheating and curing of plastics.
3. Processing of rubber and synthetic materials.
4. Sterilization of foods and medical supplies.
5. Drying and heat treatment of textiles such as rayon yarn.
6. Processing of chemicals during manufacture.

An important wartime application was the heating and gluing of plywood panels. A stack of plywood panels 4 feet by 8 feet was placed between the plates of a large hydraulic press. Three conductors or plates are interspaced with the plywood panels (negative plate at top and bottom and positive at center), and the terminals are connected to a high-frequency generator. Upon application of the voltage, every molecule of the plywood becomes stressed by the a-c field and must reverse at the rate of two million to twenty million times per second. With the watts input $W = KE^2f$, it can be understood how 15,000 to 18,000 volts at 2 megacycles or 3000 to 4000 volts at 15 to 20 megacycles will generate the necessary heat. This process of dielectric heating will bring the central parts of the plywood to the desired temperature in 20 to 30 minutes, whereas the older method using steam heat required 6 to 8 hours.

Other applications of dielectric heating concern preheating plastics for molds or curing the plastic within the mold itself. Ceramics may be heated quickly by the dielectric process.

A novel application of dielectric heating is known as the "electronic sewing machine." This machine which resembles a conventional sewing machine replaces the customary needle and thread with a pair of rotating electrodes between which a high-frequency voltage is applied. The machine is designed to make seams in thermoplastic sheet materials such as those known under the trade names of Pliofilm, Vinylite, Koroseal, and Saran. The thermoplastic sheets can be bonded when the parts to be united are raised to the proper temperature and then submitted to light pressure. The thermoplastic materials are used for raincoats, packaging of products, and other purposes. The electronic sewing machine has definite advantages where a waterproof or vacuum-tight seal is needed. Other methods of fabricating the thermoplastic sheets are (1) sewing with needle and thread, (2) bonding by hot rollers, and (3) use of solvent and pressure. The first method is weak due to the presence of holes, and it is neither waterproof nor vacuum-tight. The second method permits some flow of plastic material since heat is applied from outside; it thus results in a poorer appearance and a weaker bond. Both the second and third methods give a waterproof bond but result in some weakening of the material due to the action of the solvent.

Dielectric heating is being tried for numerous applications in the processing of foods and should have many economical uses. Some of these applications are the "popping" of cereals, the thawing of frozen products (both raw and precooked), the cooking of foods, the killing of molds in bread, and the treatment of dairy products.

Diathermy as applied to man and animals for the treatment of pain and disease is a well-known application of dielectric heating. For this application rubber-insulated coils of a conductor are substituted for the parallel plates of the dielectric heater.

Calculation for Dielectric Heating. For an illustrative problem, assume that eight 24 by 24 by ¼-inch maple panels are being glued together in a stack 2 inches high. It is desired to raise the temperature of the stack from 70 degrees F to 200 degrees F in 5 minutes, using a frequency of 1 megacycle. Calculate the power required at the load and the voltage to be applied on the assumption of a power loss of 20 per cent.

The thermal power from equation 8 is

$$P_T = 17.6Mc\,\Delta T \text{ watts}$$

$$\text{Volume of maple} = 24 \times 24 \times 2 = 1150 \text{ cubic inches}$$

$$\text{Weight of maple from Table 2} = 1150 \times 0.022$$

$$= 25.3 \text{ pounds}$$

$$M \text{ (pounds heated per minute)} = \frac{25.3}{5} = 5.06 \text{ pounds}$$

$$c \text{ (Table 2)} = 0.42 \text{ Btu per pound per degree F}$$

$$\Delta T = 200 - 70 = 130 \text{ degrees F}$$

$$P_T = 17.6 \times 5.06 \times 0.42 \times 130 = 4870 \text{ watts}$$

Adding 20 per cent for losses 974 watts

5844 watts

From equation 22 and Table 2

$$c = \frac{2248AK}{10^{10}d}$$

$$= \frac{2248 \times 24 \times 24 \times 4.4}{10^{10} \times 2}$$

$$= 2.85 \times 10^{-4} \text{ microfarad}$$

From equation 21

$$w = 2\pi fCE^2 10^{-6} \times pf \text{ watts}$$

$$5844 = 6.28 \times 10^6 \times 2.85 \times 10^{-4} \times E^2 \times 10^{-6} \times 0.0341$$

$$E = 9750 \text{ volts}$$

The required voltage is less than 15,000 and the spacing between electrodes is less than one-eighth of the wavelength.

Equipment for High-Frequency Heating. Several types of conversion units are available for producing the high-frequency power for the heating processes considered in the preceding sections. The fre-

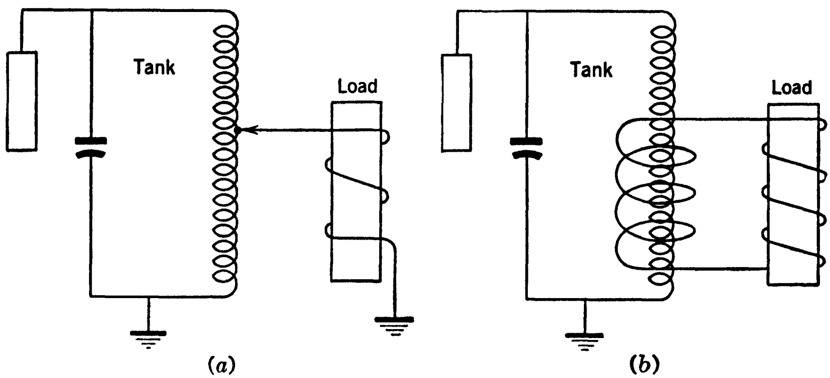


Fig. 15. Circuits for coupling an induction heating load to oscillator.

quencies obtainable and the approximate conversion efficiency may be listed as follows:

1. Mercury-arc converters.
   Frequency up to 1500 cycles.
   Approximate conversion efficiency 90 per cent.
2. Motor-generator sets.
   Frequency up to 10,000 cycles.
   Approximate conversion efficiency 65 to 75 per cent.
3. Spark-gap converters.
   Frequency 30,000 to 40,000 cycles.
   Approximate conversion efficiency 50 to 60 per cent.
4. Vacuum-tube oscillators.
   Frequency 100,000 to 100,000,000 cycles.
   Approximate conversion efficiency 50 to 60 per cent.

The units listed under 1, 2, and 3 are used exclusively for induction heating. The vacuum-tube oscillator is the only one used for dielectric heating and it is also used for induction heating where the higher frequencies are required. The mercury-arc converter may be used for melting and heating large charges where a frequency of 1000 to 1500 cycles is satisfactory. Motor-generator sets are available for various fixed frequencies up to 10,000 cycles and for inductive-heating loads. Power output of these units ranges from 7.5 to 250 kilowatts with ini-

tial and maintenance costs lower than such costs for vacuum-tube converters.

Spark-gap converters are available in units for power output from 1 to about 35 kilowatts. They are usually rated in input kilovolt-amperes and not in output kilowatts. In the radio-frequency field they have a slightly lower initial cost per kilowatt output than the vacuum-tube converter and the advantage of relative simplicity. The sparking across the gaps causes corrosion and periodic readjustment of the gaps is necessary.

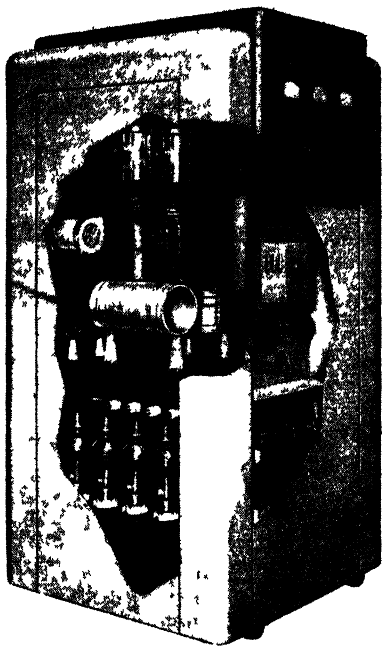


FIG 16 Construction of a high-frequency induction heating unit. (Courtesy Allis-Chalmers Manufacturing Company.)

The vacuum-tube oscillator unit for high-frequency heating may employ any of the oscillator circuits described in Chapter X. The Colpitts and the tuned-plate circuits are the ones generally employed. The oscillator unit may be connected to the load or charge in a number of ways. Thus, for inductive heating, the inductance of the tank coil may be tapped to supply the inductor as shown in Fig. 15*a*, or it may be coupled through transformer action as in part *b* of the same figure. For dielectric heating the load or charge capacitance may be substituted for the capacitor of the tank circuit. Since relatively large currents are produced by the oscillator unit, means must be employed to absorb the heat losses within the oscillator circuit. Thus the vacuum tubes employed are usually water cooled and the inductance and inductor coils are copper tubing through which water may be circulated to remove the heat loss. An oscillator unit for high-frequency heating is illustrated in Fig. 16. The tubes used for rectification are of the hot-cathode, mercury-vapor type because they are rugged and can supply the high currents required with little internal voltage drop. The oscillator tubes must be of the high-vacuum type.

All manufacturers incorporate protective devices in the oscillator type of converters to prevent damage to tubes in the event of overload, failure of water supply, and so forth. Some of these devices and their functions are as follows.

1. Thermal Element. A bimetallic element connected in the water-return line of the oscillator tube which will shut down the unit when the water reaches too high a temperature owing to insufficient flow or overload.

2. Pressure Switch. A bellow-type switch which prevents starting of the unit or opens the switch when the water pressure falls below 40 pounds.

3. D-C Overload Relay. A magnetic solenoid-type relay connected in the negative line of the high-voltage plate supply opens on short circuit, flashover, or overload of the oscillator tube.

4. Time-Delay Circuit. A circuit to prevent the application of the plate voltage until the cathodes of the rectifier tubes are brought to the proper temperature for emission.

Radio interference from high-frequency generators can be prevented by operation at frequencies differing from all communication channels or by complete shielding of all high-frequency equipment including the rooms where work is being performed.

TABLE 2. PROPERTIES OF NONCONDUCTORS

Material	Volume Resistivity (ohm-inches)	Specific Heat (Btu per pound per degree F)	Weight (pounds per cubic inch)	Dielectric Constant		Power Factor		
				1 Mc	30 Mc	1 Mc	10 Mc	30 Mc
Fiber, commercial	2×10^9	0.25	0.038	5 0		0.05		
Glass								
Pyrex "radio"	5×10^{13}	0.20	0.077	4 0	4 0	0.00075	0.001	0.001
Nylon	4×10^{12}	0.55	0.014	3.0	3.0	0.02	0.02	0.02
Phenolic insulation								
Nema grade CE	4×10^{11}	0.42	0.049	5.3	5.0	0.05	0.06	0.066
Resins								
Beetle		0.40	0.054	5.5	5.2	0.027		0.05
Rubber								
Natural gum	10^{13}–10^{15}	0.33	0.041	2.4	2.4	0.002		0.004
Maple	10^{10}	0.42	0.022	4.4		0.0341		
Mahogany	2×10^{13}	0.42	0.024	2.1	2.1	0.03	0.04	0.04

Problems

1. The surface of a round brass bar of 1.0-inch diameter is to be heated from 100 to 600 degrees F over a length of 3 inches in 1 minute. The inductor is wound of ¼-inch outside-diameter copper tubing with an inside diameter of 1.5 inches and heated by 450,000-cycle power. Neglecting inductor I^2R loss, calculate the depth of penetration, thermal power, power density, coil current, and total power delivered into the load.

2. A round hard-tempered steel shaft 1.5 inches in diameter is to be heated for a length of 8 inches. The temperature is to be raised from 70 to 1100 degrees F in 1.5 minutes using 450,000 cycles. Assume that the inductor is wound with ⅜-inch outside-diameter copper tubing to an average diameter of 2.375 inches and that the effective a-c resistance of tubing is 498×10^{-6} ohm per inch.

Calculate the depth of heat penetration, thermal power, power density, peak magnetizing force, coil current, voltage required, and total power of load.

3. A sheet of commercial fiber 10 by 20 by 0.3 inches is to be heated by the dielectric process using 1-megacycle power and to have its temperature raised 200 degrees F in 3 minutes. Assuming electrodes of same area as the fiber and a 10 per cent loss of power, calculate the thermal power required, the capacity of the load, and the voltage required.

4. A block of natural gum rubber 1 foot square and 2 inches thick is to be heated by dielectric heat using 30 megacycles of power. If the temperature is to be raised 200 degrees F in 3 minutes by electrodes of same area as load, calculate the thermal power, voltage required, and total power input to the load, allowing for 20 per cent loss.

REFERENCES

Mortenson, S. H., "Induction Heat Works Metal Magic," and H. F. Storm, "Mercury-Arc Converter Masters Metals in Deep Heating," *Allis-Chalmers Electrical Review*, March 1945.

Jordan, J. P., "The Theory and Practice of Industrial Electronic Heating," *Gen. Elec. Rev.*, December 1943.

Bendz, W. I., and C. A. Scarlott, *Electronics for Industry*, John Wiley & Sons, 1947.

Chapter XIV

RESISTANCE WELDING

Resistance Welding. Resistance welding is produced by a concentrated electric current flowing across the contact resistance between two metallic pieces held under pressure. The heating effect is represented by the equation I^2RT where I is the current in amperes, R is the contact and load resistance, and T is the time during which the current flows. In this process, the resistance R is a very important factor

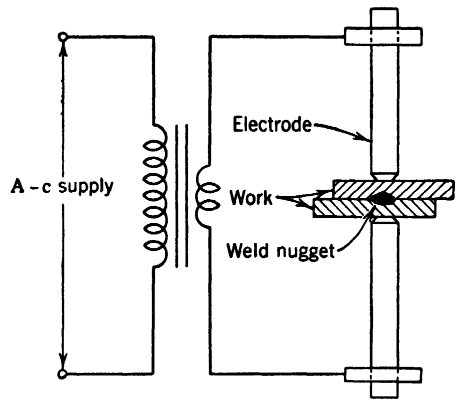


FIG. 1. Schematic circuit for resistance welding.

which is determined by nature of material, surface condition of material, and the pressure employed. The process is classified as pressure welding by the American Welding Society.

The principle of resistance welding was patented by Professor Elihu Thompson in 1886. The application of the principle is indicated in Fig. 1. Two pieces of metal are held under pressure between two electrodes having high conductivity, while a heavy a-c current from the secondary of a transformer is passed through the small contact area between the pieces. Since most of the resistance in the secondary circuit exists at the contact between the pieces in the immediately sur-

rounding metal, the energy developed in this circuit is transformed into heat I^2RT at and close to the point of contact. The metal at the point of contact quickly reaches the fusion point and the pressure of the electrodes brings about a molecular union which on cooling results in a welded joint. The voltage employed in the secondary circuit is relatively low, of the order of 2 to 10 volts. The current producing the weld is relatively large but depends on the thickness and size of the

Metals	Aluminum	Ascoloy	Brass	Copper	Galvanized iron	Iron	Lead	Monel	Nickel	Nichrome	Tin plate	Zinc	Phos. bronze	Nickel silver
Aluminum	●										●	●		
Ascoloy		●	●	●	●	●		●	●	●	●		●	●
Brass		●	●	●	●	●		●	●	●	●	●	●	●
Copper		●	●	●	●	●		●	●	●	●	●	●	●
Galvanized iron		●	●	●	●	●	●	●	●	●	●		●	●
Iron		●	●	●	●	●		●	●	●	●		●	●
Lead					●		●				●	●		●
Monel		●	●	●	●	●		●	●	●	●		●	●
Nickel		●	●	●	●	●		●	●	●	●		●	●
Nichrome		●	●	●	●	●		●	●	●	●		●	●
Tin plate	●	●	●	●	●	●	●	●	●	●	●		●	●
Zinc	●		●	●			●					●		
Phos. bronze		●	●	●	●	●		●	●	●	●		●	●
Nickel silver		●	●	●	●	●	●	●	●	●	●		●	●

FIG. 2. Combinations of materials that may be welded satisfactorily. (Courtesy General Electric Company.)

parts to be welded, the area of contact, and the material being fabricated. For welding small wires the current may be less than 100 amperes, while in fabricating heavy structural-steel parts the current may reach hundreds of thousands of amperes. Most applications of resistance welding employ alternating current and the power factor averages approximately 0.5. In special cases the power factor may be as low as 0.1 or as high as 0.9. Some modern welders employ the energy storage system wherein welding is performed by a single pulse of current. During the first quarter of this century resistance welding was confined largely to the fabrication of the ferrous metals, but with the improvement in techniques of the method, and particularly with the introduction of electronic controls covered in later sections of this chapter, it has become possible to weld successfully nearly all metals

and alloys. A chart showing the combinations of metals that may be welded satisfactorily is given in Fig. 2.

Divisions of Resistance Welding. Resistance welding may be subdivided into (1) spot, (2) projection, (3) seam, and (4) butt welding as shown in the chart of Fig. 3. These forms of welding all follow the principle previously outlined, differing only in the specific method of

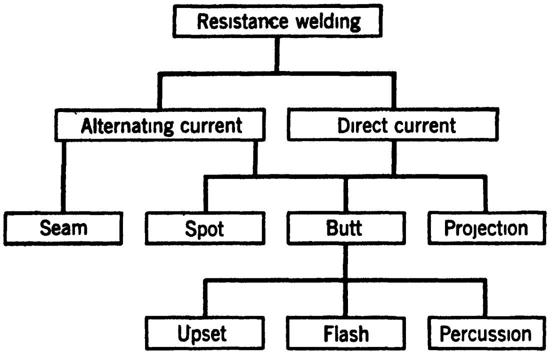


FIG. 3. Chart of the divisions of resistance welding.

application. Spot welding is the most common form. Here the welds are made by localizing the welding current through a small area in the parts being welded by the use of points or electrodes of proper shape as suggested in Fig. 1. This method is used for fabricating sheet-metal structures where mechanical strength only is required and not an air- or watertight seal. This includes such structures as metal furniture,

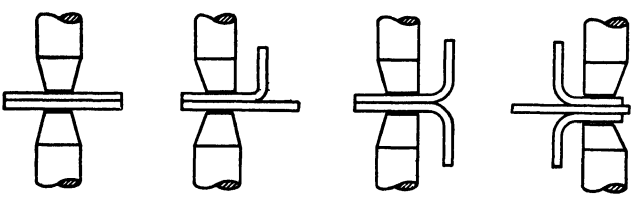


FIG. 4 Forms of sheet-metal joints.

metal cases, automobile body work, toys, refrigerator parts, and kitchen utensils. Some common forms of sheet-metal joints suitable for the spot weld process are shown in Fig. 4 and a modern heavy-duty spot welder is illustrated in Fig. 5. The time during which the current is applied is a very important element in satisfactory spot welding. As an example, when stainless steel is spot welded, the metal must be

heated for a minimum period of time to prevent carbide precipitation. Usually this time should not exceed 3 cycles on 60-cycle current, or $\frac{1}{20}$ of a second. If the heating time varies 1 cycle, the heating effect would vary $33\frac{1}{3}$ per cent, which is usually fatal to the quality of the weld. A photomicrograph of a weld of stainless steel 0.037 inch thick, made in 1 cycle ($\frac{1}{60}$ second), is shown in Fig. 6. Copper is one of the most difficult metals to weld. It can be welded and its weldability can be increased if it has a tinned surface.

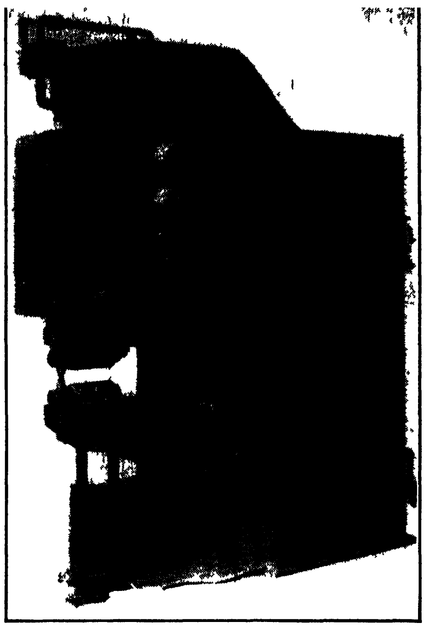


FIG. 5. Air-operated heavy-duty press-type spot welder for general purpose use with built-in electronic control. (Courtesy General Electric Company.)

A recent modification of spot welding, known as *pulsation spot welding,* makes practical the welding of two pieces of material 1 inch thick and sandwich combinations such as a $\frac{1}{2}$-inch plate between two thicker plates. The procedure for pulsation spot welding is to employ a spot-welding press of high capacity, capable of applying heavy pressure, and then through the use of electronic control to apply the power intermittently. For example, the power may be applied for 20 cycles "on" and 10 cycles "off" for ten impulses to weld two pieces of 1-inch stock. The "off" time in this process permits the water-cooled electrodes to retain their temperature below the point where they are upset excessively.

Projection welding is a modification of spot welding in which current is localized by the use of projections on the work parts instead of the use of electrodes having a small diameter or contact area. Where projection welding can be employed it is preferable to regular spot welding because the electrode maintenance is much lower and several welds may be made simultaneously by using large flat electrodes. The projection used on the work should have a rounded or dome shape as shown in Fig. 7, parts *a* and *b*, so that the weld will start with a point

contact and have a continuously increasing cross section to finish the weld. The projections should be made on the heavier piece when materials of different thickness are being welded. Silver contacts are welded

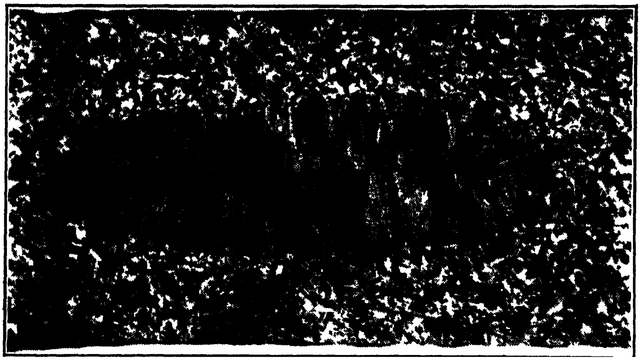


FIG. 6. Photomicrograph of spot-welded joint of two pieces of 18-8 stainless steel, 0.037 inch thick. Heat zone approximately one-half way through the weld made in 1/60 second, 50 times. (Courtesy General Electric Company.)

to various mountings by the projection method shown in Fig. 8. These contacts are prepared for welding either by forming a projection directly on the back of the silver (Fig. 7c) or by using silver steel contacts with the steel back forming a dome. Projection welding should be performed in a time period of 5 to 30 cycles.

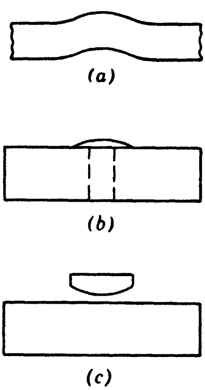


FIG. 7. Forms of projections for projection welding.

Seam welding consists of a series of overlapping spot welds. Such welds could be made on a spot welder but are produced more conveniently by two wheel electrodes, or by one wheel and a bar electrode with the work moving between them. On some seam welders the wheels are driven; on others, they are idle. A typical seam welder is illustrated in Fig. 9. The current for making seam welds is interrupted so that the metal may chill under pressure, giving a weld which simulates a continuous spot weld. In this way a pressure-tight seam weld is made. The current interruption requires the use of an electronic control to give the same "on" and "off" time to each interruption and weld. Current interruptions may be as high as 450 per minute. Seam welds are usually made under water to keep the heating of the work and welding wheels to a mini-

mum. Such cooling reduces both the distortion of the work and the maintenance cost of the welding wheels. Seam welding is used in the fabrication of oil tanks, transformers, refrigerator evaporators, gas tanks, steam radiators, all-metal vacuum bottles, and various other products.

Butt welding is a process in which two pieces of metal stock or tubes are placed end to end and welded into one piece. In the resistance-welding process two types of butt welding are employed—upset and flash. In the upset type the parts to be welded are clamped in a welding machine (see Fig. 10) in the proper alignment and are brought together with the required amount of pressure. The current is applied, heating the metal by the resistance method until the welding temperature is reached. At this point a pressure is applied which forges the parts together. The required pressure is applied manually by lever and toggle, or automatically by some hydraulic or spring system. Upset butt welding is used principally on nonferrous materials for welding bars, rods, and tubing.

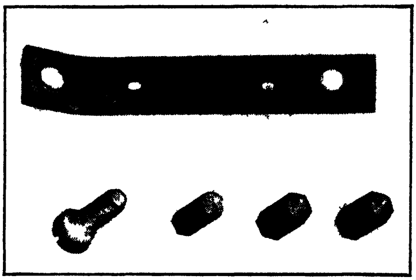


FIG. 8. Projection-welded silver contacts. (Courtesy General Electric Company.)

In flash butt welding, the parts to be welded are clamped as in the upset process and the procedure is similar except that the circuit is closed before the parts are brought together. As the parts move into contact at the correct speed, an arc is established which continues as long as this speed continues. This arc burns away a portion of the material from each piece until the welding temperature is reached. At this point the speed of travel is increased, the power switched off, and the weld upset.

Flash welding is used on steel and other ferrous alloys because it gives a better weld than the upset process. The superior weld results because the surfaces to be welded are burned away, producing a weld of clean virgin metal. Flash welding is used extensively in automobile construction on body, axles, wheels, frame, and other parts. It is also used for welding motor frames, transformer tanks, sheet metal containers, and it is applicable for steel structures of many kinds. Flash welds may be made on circular sections as small as 0.010-inch diameter and on flat sections of an area of 60 square inches.

Percussion welding is a form of resistance butt welding which uses the discharge from a capacitor for the source of energy. This process is used in factories for welding lead-in wires for incandescent lamps and electron tubes at the rate of 200 per minute. The actual duration of the capacitor discharge is approximately $\frac{1}{3000}$ of a second.

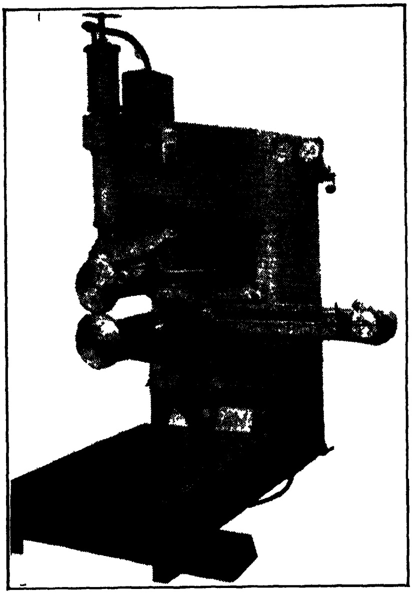


FIG. 9. Special seam-welding machine for both circular and longitudinal seam welding. (Courtesy General Electric Company.)

Miscellaneous Application of Resistance Heating. The principle of resistance heating may be employed for other heating processes such as annealing, punching holes, soldering, and brazing. Spot welders may be used for spot annealing. Butt welders are used for annealing hardened shafts and may be used for straightening tubing and other structural shapes. They may be used also to heat pipes and angles for bending. By using a tungsten punch and a copper die in spot welders, holes can be burned in hardened steel parts such as saws, thin milling cutters, and flat spring stock. These holes are not so accurate as drilled stock but they are clean, they do not start cracks in the adjacent material, and the annealing action does not extend far from the hole. The use of carbon, graphite, or tungsten as the top electrode in a welder permits soldering and brazing operations. Here the heat is produced in the electrode rather than in the work. Portable tongs with carbon electrodes utilizing the resistance method are satisfactory for brazing electrical connections in transformers, motors, and other types of electrical equipment. Through the use of silver solder a stronger joint and one that will withstand a higher temperature is obtained at a lower cost.

Welding Equipment. Welding machines must perform the following functions in the resistance-welding process.

1. Support the work and apply the pressure or squeeze.
2. Pass current for the weld.

3. Hold (with current off) for the weld to set.
4. Release work—off.

The work is held and current is applied through the use of electrodes which may be points, rods, wheels, or flat surfaces. These electrodes should have high conductivity (except for soldering and brazing). Originally only copper was used for electrodes. It has the disad-

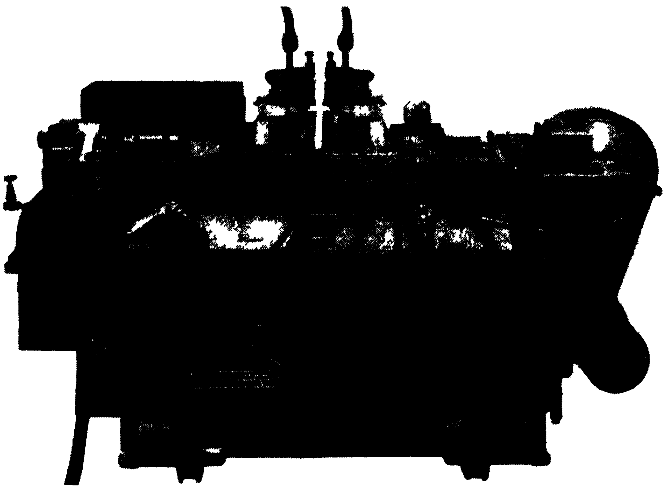


FIG. 10. Butt welder, 200 kilovolt-amperes. (Courtesy General Electric Company.)

vantage of (1) low compressive strength, and (2) a low annealing temperature causing it to misform very rapidly. Several types of electrode materials have been developed that are more suitable than copper. Many alloys of copper with cadmium and other elements have been developed with varying degrees of conductivity and strength so that satisfactory electrodes for welding all classes of materials are available. In general, low-conductivity material should be welded with a relatively high-conductivity electrode; high-conductivity materials with low-conductivity electrodes.

The pressure on the work before and during the welding process is produced manually by foot pressure, by spring-mounted electrodes, or by hydraulic operation. Air pressure in piston-type chambers is frequently employed. Pressure may be actuated by electromagnetic solenoids, or by motor-operated cams.

The current supply to the work is furnished by a transformer (except for percussion welding) which is built as an integral part of the welding machine.

The control for the various functions in the welding process is the key to satisfactory resistance welding and will be treated in the remainder of this chapter.

Steps in Welding Control. The process of making a typical weld can be divided into the four steps illustrated in the welding time and duty cycle of Fig. 11. First, the work pieces must be pressed together or squeezed in preparation for the welding current. Second, the welding circuit should be closed and the proper value of current applied for the correct interval of time. Third, the current should be turned off and the work held for a suitable time interval to permit the weld to set. Fourth, the pressure on electrodes should be released and the work removed during the "off" period. The important factors in this welding cycle are: (1) the pressure on work, (2) the current that flows, and (3) the time that the current flows.

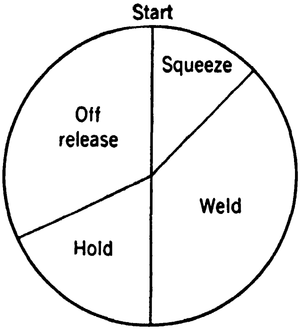


FIG. 11. Typical duty cycle of a resistance welding operation.

Methods of Control. The four steps in a resistance welding operation outlined in the preceding paragraph may be performed (1) by manual operation, (2) by semiautomatic means, or (3) by fully automatic equipment. The duty cycle on the early spot and butt welders was controlled manually and mechanically. On the spot welder a foot-operated lever was designed so that the first part of a downstroke closed the electrodes on the work and applied spring pressure (squeeze), the second continuing part of the stroke closed the circuit of the primary of the transformer (weld), the third part of the downstroke opened the electric circuit but maintained pressure on the work (hold), and the fourth or upstroke of the foot lever released the pressure and the work, thus completing the welding cycle. In the semiautomatic method of control, some one or two of the factors such as pressure, circuit closing, current magnitude, or timing may be performed automatically while other steps in the cycle are controlled by the operator. For example, the pressure on the work may be produced by a motor-controlled or magnet-controlled spring, by a solenoid-operated piston in a cylinder under air, or by hydraulic pressure. The circuit may be closed by a magnetic contactor, a motor and cam action, or an electronic switch. Similarly the timing may be accomplished by a motor-

operated timing device, a magnetic delay device, or an electronic timing circuit. In the fully automatic method of control, all steps in the complete duty cycle are performed automatically in sequence leaving nothing to the operator except the initiation of the cycle through push-button control. Under the automatic method the individual steps may

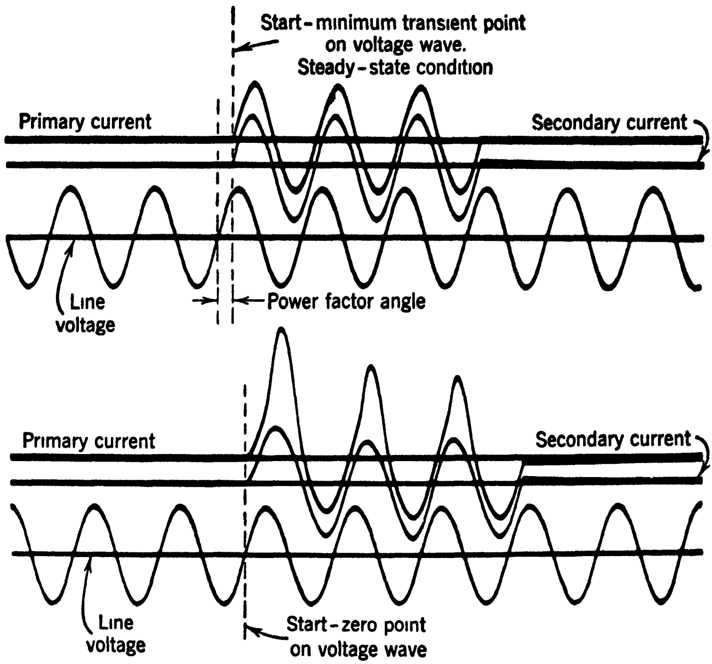


FIG. 12. Oscillogram illustrating the power factor angle. (Courtesy General Electric Company)

be mechanical or electronic though nearly all present-day welders use electronic controls.

The magnitude of the current used in the welding process may be controlled by varying the voltage applied to the welding transformer or by varying the tap connection on the primary of the transformer. These methods give rather coarse steps of variation. A smooth and stepless control can be obtained by a phase-shift timing for firing the electron trigger tubes in the primary circuit of the welding transformer. This method will be covered in a later article on heat control.

The magnitude of the heat, developed in the load by the secondary current of the welding transformer, depends upon the instant in the

impressed voltage cycle that the primary circuit is closed. This follows because for some conditions a transient current is produced in the secondary which increases the magnitude of the resultant current for one or more cycles. Accordingly the amount of heat will vary even though the time of current flow measured in cycles remains the same. This variation in the amount of heat is important in welding materials that are critical as to welding temperature. Also the transient current

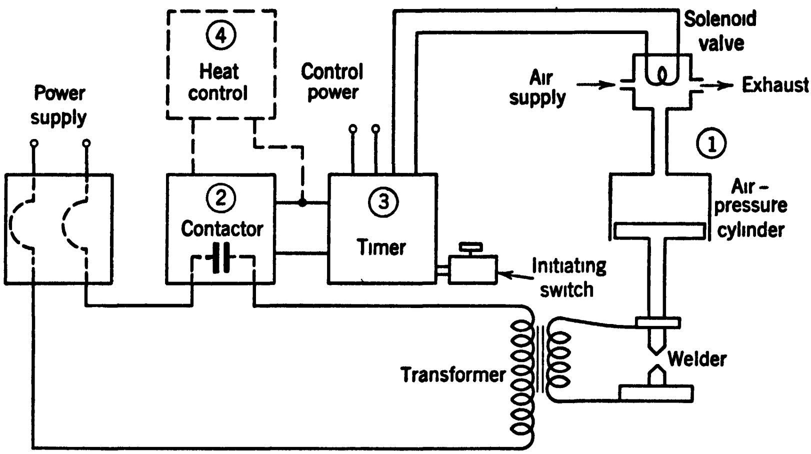


FIG. 13. Block diagram of electronic controls applied to welding processes

surges tend to burn electrodes and increase the maintenance cost on the welder.

The phenomenon of transient current suggested above is illustrated in Fig. 12. These oscillograms show that the transient can be avoided by closing the primary circuit at the power factor angle. An exact timing of the circuit can be obtained by using a phase-shift control on a thyratron to fire an electronic switch. The use of this method is known as *synchronous timing.* Synchronous timing requires special electronic controls and an additional cost of equipment which is justified in many applications. All mechanical circuit-closing devices and some electronic controls give nonsynchronous switching (i.e., the heat energy developed may vary with different welds).

Block Diagram of Control. The various functions in the welding cycle are usually performed by different units of equipment. These units and their interrelationship are illustrated in the block diagram of Fig. 13. Pressure is applied to the work by some mechanical unit shown at (1). The circuit for the primary of the transformer is closed

by the contactor (2). The time that the circuit is closed is controlled by a timer (3), and, if special control of the magnitude of the current for the heating is necessary, it may be provided by another unit known as the heat control (4). It will be noted that the timer may act as a

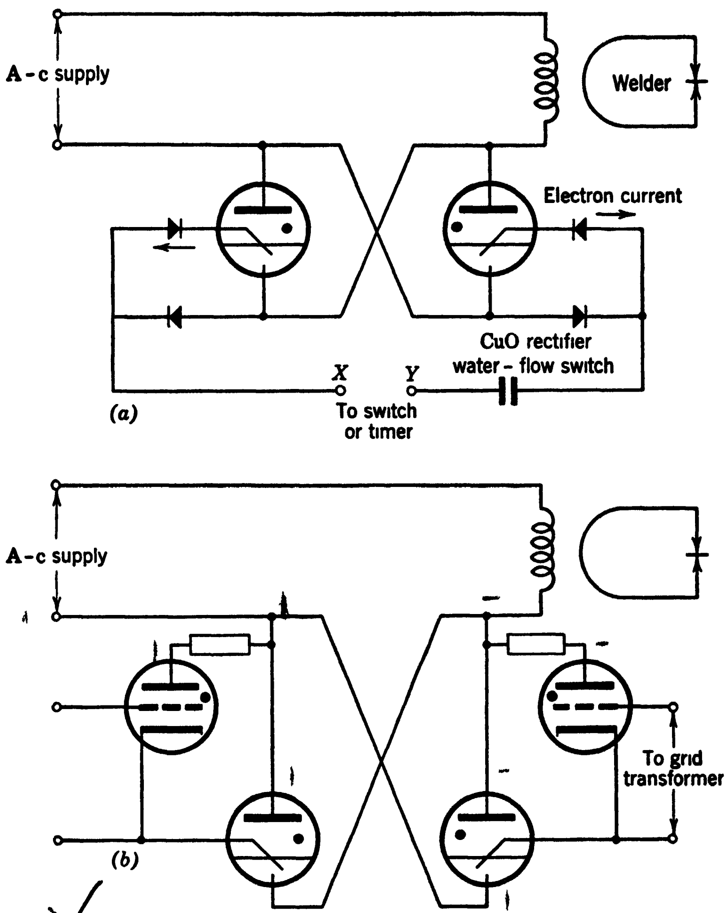


FIG. 14. Inverse parallel trigger circuits for contactors in welding control.

master control for the function of the other units. A welding installation may employ all or a part of the individual units shown in Fig. 13, depending on the refinement needed in the welding process.

Electronic Contactors. Electronic contactors consist of two trigger tubes connected in an inverse parallel arrangement to function as a single-pole switch. Either thyratrons or ignitrons may be employed.

Most welders use ignitrons as shown in Fig. 14*a*. Here the ignitrons are excited by copper oxide (or Rectox) rectifiers. The arrows indicate the direction of electron flow and, with the gap from X to Y closed by some switching arrangement, the operation is obvious (see page 236 for ignitron firing). A commercial electronic contactor (Weld-O-Trol) using this circuit is shown in Fig. 15 with the rectifiers appearing on

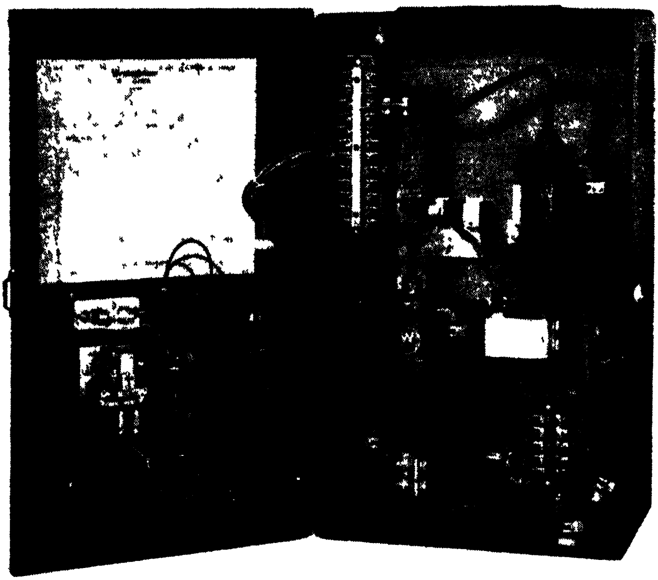


Fig. 15. The Weld-O-Trol—an electronic switch using ignitron and copper oxide rectifiers for welding. (Courtesy Westinghouse Electric Corporation.)

the left and the water-cooled ignitrons on the right. A thermostatic device in the water-circulating system controls the water-flow switch (Fig. 14*a*) so that the ignitrons are protected (shut off) from overheating which may result from a stoppage of water flow or inadequate circulation in the cooling system. The function of the electronic switch is to open or close the primary circuit subject to the firing and timing control which resides in a separate device connected to points X and Y. The operation of the welder may be synchronous or nonsynchronous, depending on the characteristics of the timing device.

Another type of electronic contactor (switch) for welders uses thyratrons instead of copper oxide rectifiers for firing the ignitrons (see Fig. 14*b*). Here the instant of firing the ignitrons is controlled by the voltage on the grid of the thyratron which fires the thyratron and

in turn the ignitron. This circuit gives a superior control because (1) the actual instant of firing is very precise, permitting synchronous timing, and (2) the time angle for firing may be adjusted to control the magnitude of the current loops (half-cycles) and thus control the rate of heat application. Both of these advantages are important in welding materials having critical welding characteristics. Thyratrons may be used directly for switching the power on low-capacity welders.

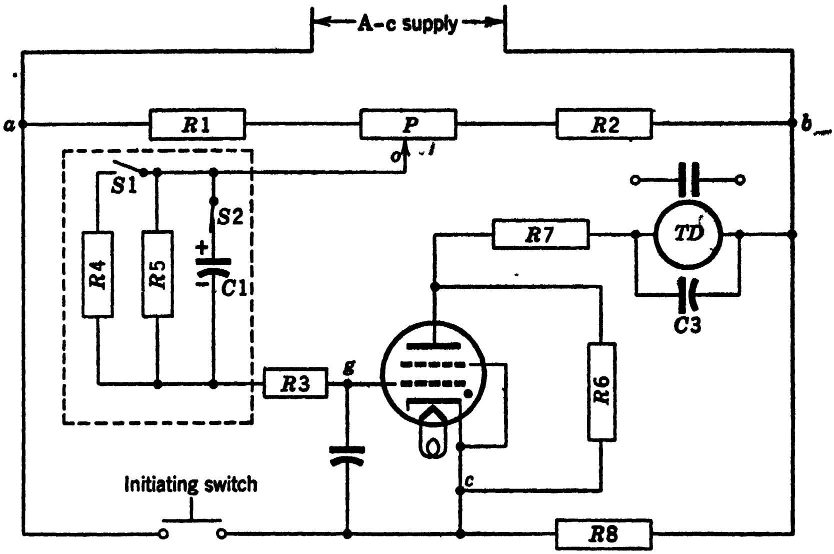


FIG. 16. Simple timer circuit using a shield-grid thyratron and a magnetic relay.

The simple switching circuit for such application is shown on page 238, Fig. 41. Obviously excellent timing and current control can be obtained through the use of suitable timing control applied to the grids of the thyratrons.

Electronic Timers. Many circuits may be used for electronic timing of the steps in the welding process. A simple and widely used circuit for measuring a definite time interval is given in Fig. 16. The key to the operation of this circuit lies in the *RC* circuit shown within the dotted rectangle. The starting of the time interval is produced by the closing of the initiating switch, and the ending of the time interval is effected by the operation of the telephone-type magnetic relay *TD*. To understand the action taking place note (1) that 115 volts a-c supply is impressed across points *a* and *b* which are the terminals of the potential divider circuit consisting of resistances *R*1, *R*2, and *P*.

The potentiometer P permits a variation of the a-c potential between points a and o and b and o. Note (2) that, as long as the initiating switch is open, both the cathode and anode of the shield-grid thyratron are connected to the right side of the a-c supply line b and hence are at approximately the same potential so that no conduction can take place. Note (3) that, as long as the initiating switch is open, the cathode and control grid g of the thyratron are connected to a circuit having an applied potential equal to o-b. Hence on the half-cycles when the cathode is negative, grid rectification takes place and electrons pass from b through $R8$, through cathode to grid, and thence through $R3$ and $R5$. After a few cycles of grid rectification the potential across the capacitor $C1$ (drop across $R5$) rises to a nearly constant value determined by the a-c potential across o-b.

The timing cycle is started by closing the initiating switch which connects the cathode of the thyratron to the supply line at a. This action permits the a-c line potential to be applied across the cathode-anode circuit (c, tube, $R7$, and relay TD). On the half-cycle, when the anode is positive the thyratron is ready to conduct (cathode to anode) but is prevented from doing so by the negative d-c grid potential which has been built up on the lower side of the timer capacitor $C1$. Now to understand what happens next, note that the closing of the initiation switch has *shifted the a-c potential applied between the cathode and grid from the value across* o-b *to the value across* a-o, that when point b is negative the cathode on the other side of $R8$ is positive instead of negative as previously indicated, and that for this condition the a-c voltage applied to the grid (across a-o) is negative with respect to the cathode so that grid rectification does not take place as before. Now the value of resistance $R1$ and $R2$ are selected so that the IR drop across $R2$ (plus part of P) will be sufficiently larger than that across a-o (plus part of P) so that the negative potential built up on the lower side of capacitor $C1$ will keep the thyratron from conducting when the initiating switch is closed. The timing action is illustrated in the graph of Fig. 17. Since the a-c voltage across a-o applied to the grid and cathode after closing the initiating switch is lower than before, condenser $C1$ discharges slowly and the d-c voltage across it decays as shown. In the meantime the new a-c component applied from a-o is superimposed in series with the d-c capacitor voltage and the resulting voltage approaches the critical firing point for the thyratron which depends on the characteristic of the tube employed. This critical point is in the region of zero resultant voltage or equal

cathode and grid potentials. When the decay of the potential across $C1$ permits the thyratron to conduct, relay TD is energized and opens (or closes) contacts which complete the timing cycle.

Capacitor $C3$ serves to hold relay TD and prevent chattering during the half-cycles when the tube does not conduct. Resistor $R7$ protects the tube from "transient" rush of current into $C3$ when the tube con-

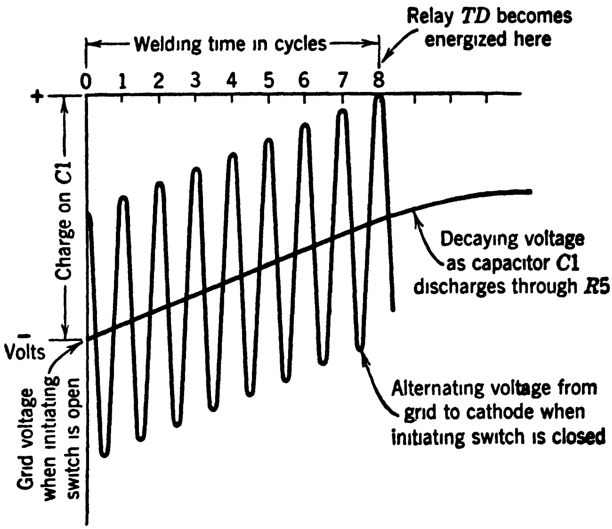


FIG. 17. Graphical illustration of the operating principle of the timing circuit of Fig. 16.

ducts. Resistor $R6$ permits a small a-c current to flow through relay TD and keep it from being sluggish in action.

The timing interval is controlled by the setting on the potentiometer which determines the initial negative potential built up on the timing capacitor $C1$. Another and larger range of timing can be secured by closing switch $S1$ which will increase the rate of decay of the d-c voltage across $C1$. The time delay can be made ineffective by opening the timing capacitor circuit at switch $S2$. The usual range of timing for this type of circuit is from 3 to 60 cycles on a 60-cycle circuit.

The timing circuit of Fig. 16 with a slight modification may be combined with an ignitron electronic switch (Fig. 14) to serve as a semiautomatic timer to measure weld time only. The timing circuit of Fig. 16 may be employed also for automatic spot welding wherein all steps in the welding cycle are controlled in the proper sequence. A complete sequence timer employs four circuits like that of Fig. 16 plus

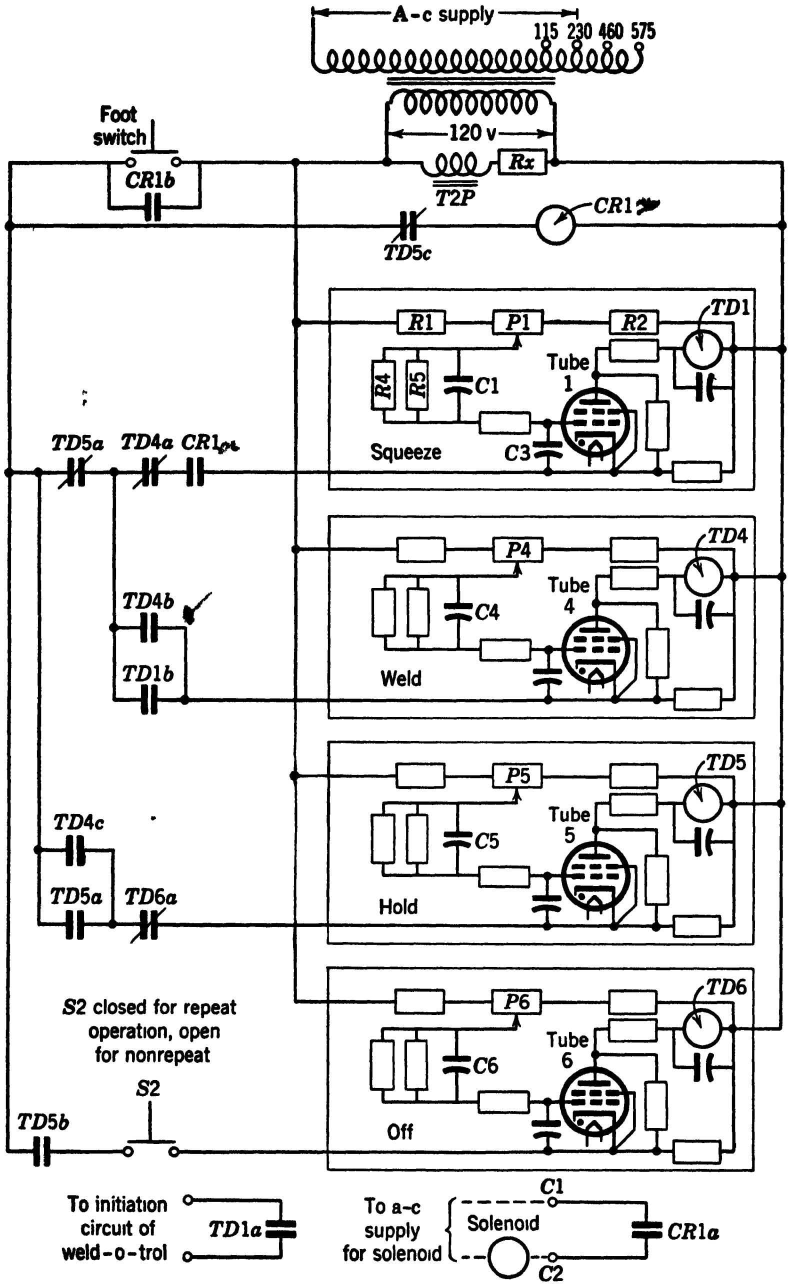


FIG. 18. Schematic circuit of a sequence timer.

relay contacts, an electronic switch, and a foot-operated initiating switch. A schematic circuit for the multiple-timer operation is given in Fig. 18 and the sequence of operation is as follows: The closing of the foot switch causes relay *CR*1 to operate and close three contacts. Contact *CR*1*b* places a shunt circuit around the foot switch, *CR*1 energizes the top (squeeze) timing circuit, contact *CR*1*a* energizes the solenoid valve (see bottom), and air pressure applies the squeeze. When the "squeeze" timing circuit completes its cycle, relay *TD*1 closes two contacts. Contact *TD*1*b* energizes the "weld" timing circuit and *TD*1*a* closes the circuit to the Weld-O-Trol (contacts *X* and *Y*, Fig. 14*a*). Energy now flows into the weld for the period determined by the "weld" timing circuit. At the close of this period relay *TD*4 operates three contacts. Contact *TD*4*b* places a shunt around *TD*1*b*, *TD*4*c* energizes the "hold" timing circuit, and the opening of contact *TD*4*a* de-energizes the "squeeze" timing circuit, which action, in turn, de-energizes relay *TD*1 and opens all its contacts. The opening of *TD*1*a* opens the power circuit through the Weld-O-Trol electronic switch. Next assume that switch *S*2 is closed for repeat operation. Now when the "hold" timing circuit completes its period, relay *TD*5 operates four contacts. Contact *TD*5*b* energizes the "off" timing circuit, *TD*5*a* (opens) de-energizes the "weld" timing circuit, and *TD*5*c* (opens) de-energizes relay *CR*1, which in turn opens *CR*1*a* controlling the solenoid pressure valve. If continuous-sequence cycles are desired the foot switch is held closed, and, when the "off" timing circuit completes its period, relay *TD*6 operates, de-energizing the "hold" timer at contact *TD*6*a*. The de-energizing of relay *TD*5 releases all its contacts which results in the complete restoration of the circuits to the initial condition ready for a repeat cycle of operation.

Precision timing requires circuits employing more tubes and the omission of magnetic relays. One important part of precision timers is the trailing-tube circuit shown in Fig. 19. This circuit, consisting of two thyratrons and three transformers, constitutes an intermediate link in a precision control circuit. In operation, tube 1 is "fired" by a timing or signal voltage applied to its grid; then, after tube 1 has conducted for a half-cycle, tube 2 is caused to "fire" and conduct for the next half-cycle. Thus, tube 2 always trails tube 1 in action. To understand this operation it should be noted that the grid of tube 2 is excited by a-c voltages from two transformers. The a-c exciting voltage from transformer *T* is 180 degrees out of phase with the a-c supply from the left. Thus, whenever the anode of tube 2 is made

positive, the a-c voltage applied to its grid from transformer T is negative which tends to hold the tube from conducting. Hence as long as tube 1 does not conduct, tube 2 will not, and the circuits to the right are held open by the inverse parallel connection.

Assume now that a timing signal is applied to tube 1 which makes its grid positive. Tube 1 will now conduct (assuming its anode is positive) for the remainder of that half-cycle and current will flow through the primary windings of both the feedback and grid-control transformers. This current will induce voltages in the secondaries of

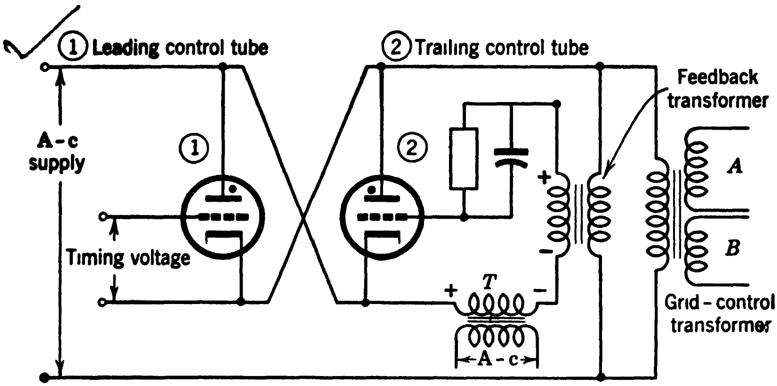


FIG. 19. Trailing-tube circuit.

both transformers. The secondary of the feedback transformer is connected so that its induced voltage "bucks" (180 degrees out of phase) the a-c voltage produced by transformer T. However, during this half-cycle the anode of tube 2 is negative and so it does not conduct. When this half-cycle ends, the applied supply voltages pass through zero and, as they rise on the next half-cycle, the anode of tube 2 becomes positive. During this transition period the primary current and the secondary voltage of the feedback transformer lags (the supply circuit) because of the leakage and other reactance in the parallel-connected primaries of the grid-control and feedback transformers. This lag of voltage in the secondary of the feedback transformer holds a resultant positive potential on the grid of tube 2 for a short interval and causes the tube to conduct for the second half-cycle. In the complete welding-control circuit, the two secondaries of the grid-control transformer, A and B, are used to excite the grids of another pair of thyratrons, which in turn are used to fire a pair of ignitrons in the electronic switch. In some welder-control circuits, the feedback trans-

former is connected across the primary of the welding transformer and thus causes tube 2 to trail tube 1.

This trailing action of tube 2 assures that any timing device employing this circuit will always give an integral number of cycles (an even number of half-cycles). This feature is important in welding operations because a weld period containing an odd number of half-

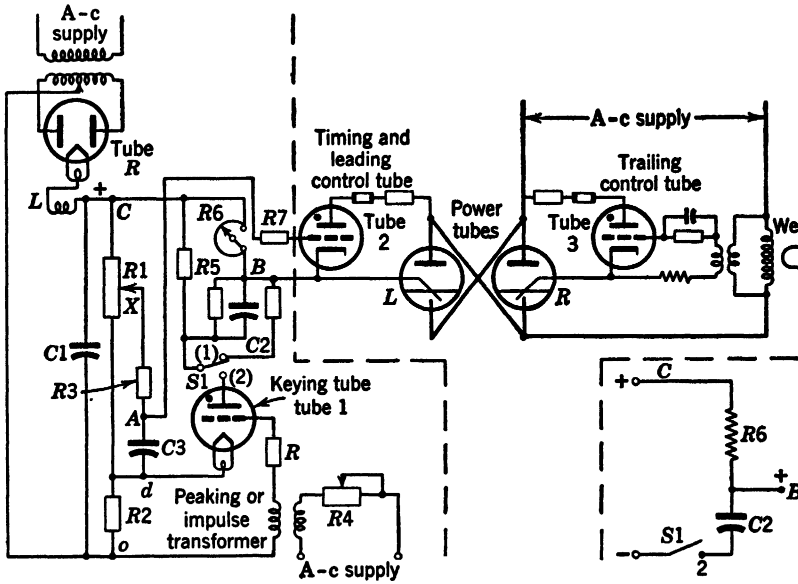


FIG. 20. Circuit of a spot-welder control without phase control.

cycles may leave residual flux in the welding transformer and result in unequal heating for different welding duty cycles.

The trailing-tube circuit of Fig. 19 can be added to Fig. 14*a* by connecting the two thyratrons 1 and 2 at points *X* and *Y* and connecting the primary of the feedback transformer across the welding transformer. A somewhat similar addition can be made to Fig. 14*b*.

The trailing-tube circuit is utilized in the spot-welder timing-control circuit of Fig. 20. Here direct current is used to produce the timing interval and three thyratrons perform the timing and firing functions. In this circuit the direct current is provided by a double diode (two-anode) rectifier at the left with reactor *L* and capacitor *C*1 serving as the filter. The filtered direct current is applied across the voltage divider formed by resistors *R*1 and *R*2. The voltage drop *o-d* across resistor *R*2 provides a negative bias for thyratron tube 1 and normally holds it from firing. Tube 1 can be fired at a very definite time in the

impressed a-c voltage wave by the peaking or impulse transformer connected in its grid circuit. The student should recall the theory of the peaking transformer explained on page 225. The time of applying heat (impulse) is controlled by an adjustment of variable resistor $R4$.

With $S1$ in position 1, the cathode-anode of tube 1 is open and the tube cannot conduct. When switch $S1$ is moved to position 2, tube 1 will conduct the instant that the next positive "peak" is induced by the peaking transformer. Thus direct current will flow from point d through tube 1, resistor $R5$, and to the point $+C$. Because of the action of a thyratron on direct current, it is obvious that a continuous undisturbed current will flow through tube 1 and $R5$ until the anode circuit is broken at some later time by moving switch $S1$ back to position 1.

The timing of the weld period and the actual firing of the left ignitron is performed by tube 2. To follow this action, note that the grid of tube 2 is held at the potential of point A and that A is at a potential between that of point d and point X on resistor $R1$. Since the potential at A is determined by the charge on the condenser $C3$ and since this charge can be varied only by current flowing through high resistance $R3$, it follows that this potential is quite stable and is influenced little by fluctuations in the supply voltage. Next observe that, with switch $S1$ in position 1, the cathode of tube 2 (point B) is at the same positive potential as point C. Thus tube 2 does not conduct because its grid potential is negative at A. Now when switch $S1$ is moved to position 2, tube 1 conducts at the next peaking voltage surge and the potential on the cathode of tube 2 switches instantly from + to − and fires. To understand this phenomenon, note that an $R6C2$ timing circuit is connected in parallel across resistor $R5$. When tube 1 conducts, d-c voltage Cd minus tube 1 arc drop appears across resistor $R5$ and in turn across $R6C2$. From the detailed $R6C2$ circuit (lower right, Fig. 20) it will be observed that, before $S1$ is closed, point B is + like C and both sides of the condenser $C2$ are at the same potential. Now when switch $S1$ is closed, a few electrons rush to the lower plate of $C2$, and simultaneously other electrons are repelled from the upper plate and instantly lower the potential of point B (cathode of tube 2). The number of electrons taking part in the action is relatively small, being limited by the resistance $R6$. However, the phenomenon instantly lowers the potential of the cathode of tube 2 below that of its grid and conduction takes place which fires ignitron tube L. On the succeeding half-cycle, tube 3 trails and fires the other ignitron tube R. Tubes 2 and 3 and the ignitrons continue to fire, while cur-

rent flowing in resistor $R6$ charges capacitor $C2$ and causes the potential on its top plate (point B) to rise. When the potential of B rises to the potential of the grid of tube 2, tube 2 ceases to conduct and trailing tube 3 stops after completing its half-cycle. The period required for capacitor $C2$ to charge up to the cutoff point depends upon

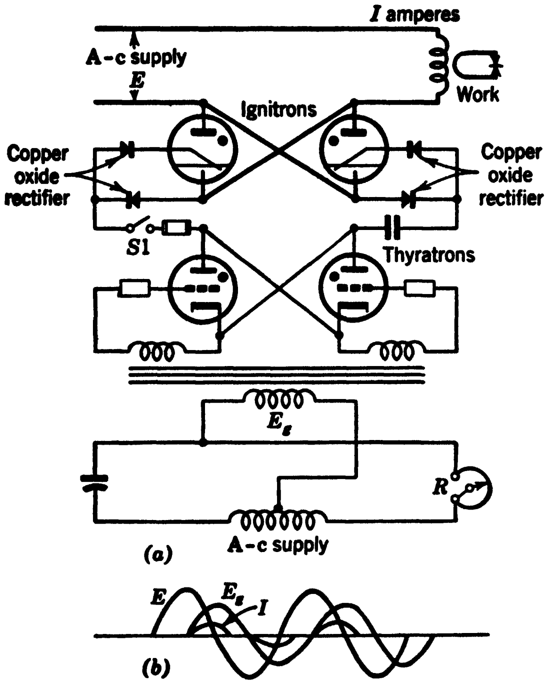


FIG. 21. Electronic contactor with simple phase control added.

the resistance of $R6$. This period is adjustable from one half-cycle to one half-second or a welding period of 1 to 30 cycles.

The control resulting from a variation of $R4$ permits firing at the power factor angle and gives sychronous timing. The adjustable point on $R1$ permits a slight variation of grid voltage to take care of variations in tube characteristics and variations in the magnitude of the resistor and capacitor components. The trailing-tube circuit assures integral cycles of weld time and still allows for small variations in the timing of the RC circuit.

Heat Control. The term heat control refers to the amount of heat energy delivered to the work in a welding operation. The total heat energy delivered is determined by the expression $H = I^2RT$. Where

R is constant, the energy may be controlled by varying the time T or by varying the current I. The preceding discussion has covered the circuits employed for controlling time or the period of application. For some metals the welding time may be varied over fairly wide

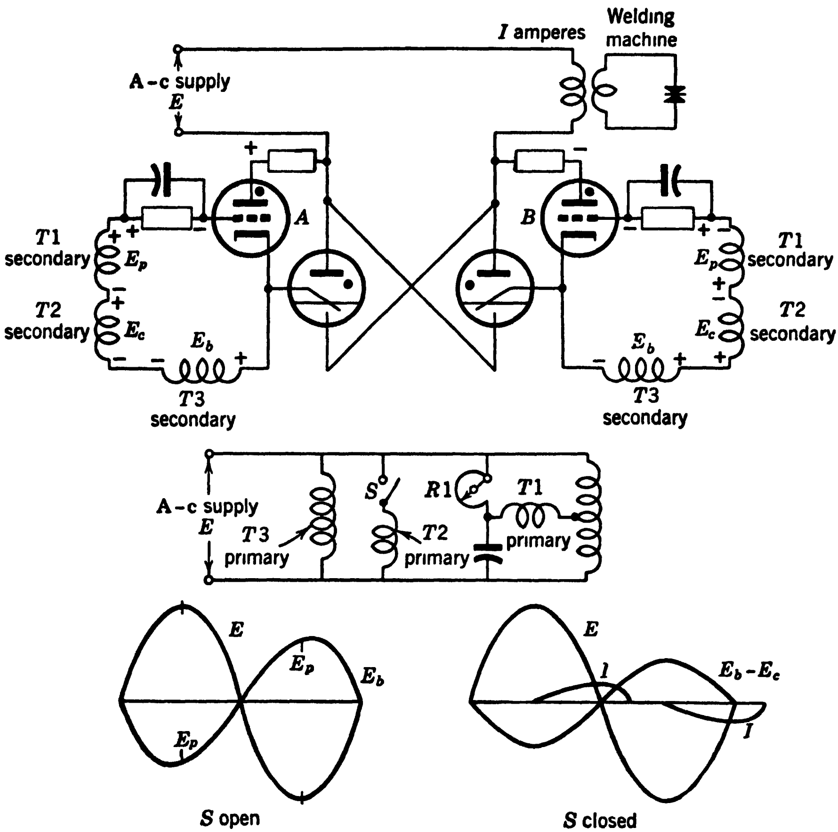


FIG. 22. Phase-shift control for resistance welding.

limits to secure the proper amount of heat energy. However, for other metals the proper welding time may be fixed by the nature of the metal itself, and here it is necessary to control the quantity of heat applied by a control of I. It was suggested earlier that the current flowing in the work may be controlled by using different taps on the primary of the welding transformer. This method gives rather large steps of current change and requires some time to effect the change. A nearly stepless control of current can be secured quickly by shifting

the phase of the exciting voltage applied to the grid of the firing thyratron. This principle was discussed on page 125 and a simple circuit for applying the principle to a welder control is given in Fig. 21. The phase shift and the firing point of the thyratrons are controlled by the dial on resistor R and may be adjusted to the power angle or a later point in the cycle. In all cases, the current loops for each half-cycle will be equal and synchronous welding will result.

A second circuit for producing a phase shift and firing of ignitrons

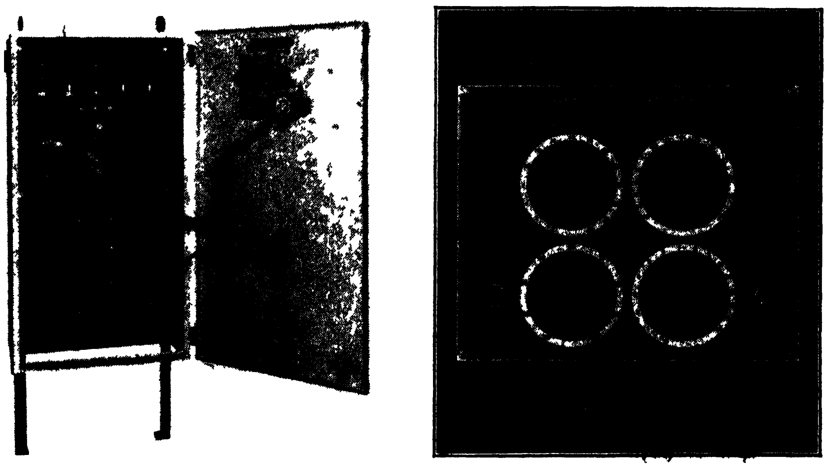


FIG 23 Thyratron spot, pulsation, and seam-welding controller with phase control. (Courtesy General Electric Company.)

is shown in Fig. 22. In this circuit three transformers are employed with their secondaries in series in the grid circuit. Transformer $T1$ is a peaking or impulse transformer which provides the peaks for firing tubes A and B and, in turn, the ignitrons. The phase of the peak voltages induced by the impulse transformer is controlled by the setting of resistor $R1$ in the bridge circuit. Transformer $T3$ alone induces a voltage E_b in the grid circuit in opposition (180 degrees out of phase) with the applied a-c E (anode circuit) which has a magnitude sufficient to prevent the thyratrons from being fired by the peaks from transformer $T1$. Transformer $T2$ induces an a-c voltage E_c which is in phase with the applied voltage E and is of such magnitude that when S is closed it bucks the secondary voltage of $T3$ and permits the thyratrons to fire. The opening of switch S permits the voltage of $T3$ to prevent firing. Thus, the period of the weld time is controlled by the operation of switch S.

An excellent welding control can be obtained by combining the features of the circuits in Figures 19, 20, and 22. Such a combination will utilize the timing control to the left of the dotted line in Fig. 20 with the grid circuit of tube 2 connected to the grid of tube 1 of Fig. 19. Then the secondaries of the grid transformer of Fig. 19 should be substituted for the secondaries of transformer *T*2 in Fig. 22.

Voltage and Current-Compensation Control. In some applications the quality of a weld may depend on factors other than that of the

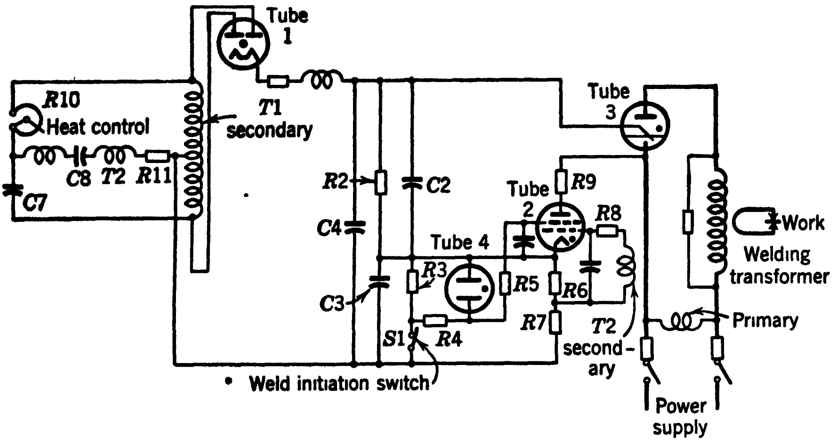


FIG. 24. Circuit for half-cycle welding having heat control.

precise timing and the proper phase shift of the firing signal. Such factors are the surface condition of the work, the pressure applied to the work, and the variation in welding-current magnitude caused by changes in welder circuit impedance, welder power factor, or power-supply voltage. The resulting variation in welding current (and heating) may require further refinement in electronic control. Such refinement is accomplished by additional circuit elements classified as voltage compensation and current-regulation compensators. These compensators utilize the phase-shift method previously explained wherein a peaked triggering voltage is supplied to the grids of the firing thyratrons. The purpose of a voltage compensator is to advance, automatically but arbitrarily, the firing point of a component of control by a definite angle so as to compensate for a given reduction in power-supply voltage. One system of voltage compensation is to employ a small "dummy" load in a part of the control circuit. The input to the dummy load varies with the supply voltage and this variation is the basis of the change in phase of the power tube firing for maintain-

ing a constant heating value in the welding current. Current compensation should be controlled by current changes in the secondary of the welding transformer. The circuits required for voltage compensation and current regulation are outside the scope of this text.

Preheating and Postheating Controls. In some welding applications it is desirable to preheat the work before applying the welding current; in others it is desirable to reduce the rate of cooling (annealing) after the welding current has ceased. These applications require an extension in the number of steps in the welding cycle. Each added step will require another timing circuit and another current-control unit in the complete sequence-control unit.

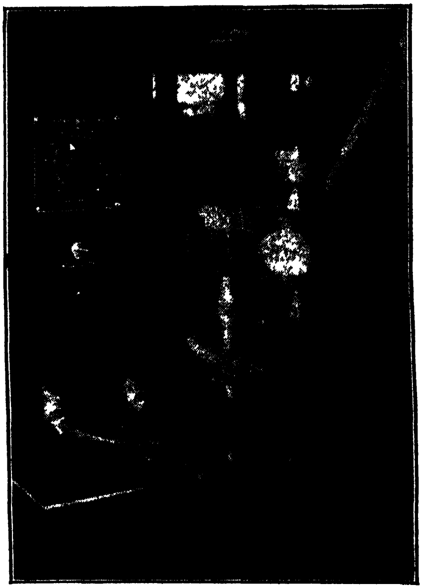


FIG. 25. Bench-type spot welder operated by half-cycle. (Courtesy General Electric Company.)

Seam and Pulsation Welding. Resistance seam welding consists of a series of successive welds between two layers of metal for the purpose of forming either a gas-tight line weld, or the equivalent of a series of spaced spot welds. Welding current is applied intermittently at definite "heat" and "cool" time intervals as the work progresses between roller-type electrodes. Successive welds are made without breaking the low-resistance secondary circuit of the welding transformer, therefore making it important that the impulses of welding current are free from starting transients. Synchronous firing and a timer giving an integral number of cycles for heating are necessary for seam welding. Circuits for seam-welding control employ the principles previously covered for spot welding. In these circuits a separate dial or potentiometer control is used for the "heat" and "cool" timing. The welding cycle is started by closing an initiating switch and is continuous with alternate heat and cool periods until stopped by opening the initiation switch.

Pulsation welding (interrupted spot welding) consists of making a spot weld by means of a spot welder but using seam-weld timing for a

definite number of current impulses. As an example, pulsation welding may use 10 cycles of "heat" time and 4 cycles of "cool" time repeated alternately for five current applications requiring an over-all welding time of 66 cycles. A combination circuit for effecting either spot-, seam-, or pulsation-welding control for a resistance welder may be built into a single panel unit as illustrated in Fig. 23.

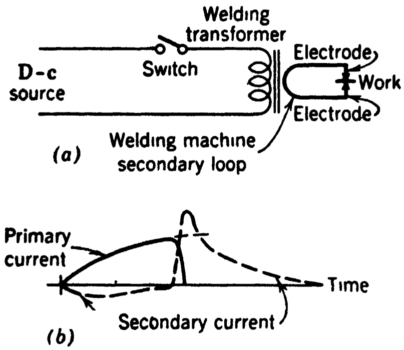


FIG. 26. Schematic circuit and curves illustrating a magnetic energy storage system for spot welding.

A semimechanical type of seam-welding control consists of a synchronous motor-driven chain having extended link pins for holding removable metallic and nonmetallic buttons. These conducting and nonconducting buttons may contact a set of brushes at the rate of one button per half-cycle or one button per cycle, depending upon the motor-gear-box ratio. The brushes render the firing tubes conductive in accordance with the timing pattern set up on the motor-driven chain. By means of auxiliary link contacts and auxiliary brushes, the timing may be initiated at a definite point on the timing chain to provide special timing patterns. Any special timing pattern can be provided by full electronic control, although there are cases where the additional complications and expense may not be justified, thus making the chain timer preferable.

Bench-Type Spot-Welder Controls. Bench-type spot welders for fabricating small parts may require either a large current for a portion of one half-cycle, or a comparatively small current for several cycles. For a large current, a half-cycle welding control using an ignitron may apply a single unidirectional impulse controlled by phase shift over an angle varying from 20 to 300 degrees. A schematic circuit for a welding control of this type is shown in Fig. 24. This circuit is of interest because it utilizes the discharge of a capacitor for firing the ignitron. A rectifier and filter on the left applies a d-c voltage across the potential divider $R2$ and $R3$ with a major part across $R2$. The discharge capacitor $C2$ charges to the voltage of $R2$. The smaller voltage drop across $R3$ and $C3$ is applied to $R4$ and the voltage-regulator tube 4 to hold the shield grid negative, and also across $R6$ and $R7$ to apply a

d-c negative bias to the control grid of tube 2. An a-c peak-firing voltage is applied to the control grid of tube 2 through its grid transformer $T2$, which voltage is large enough to overcome the negative d-c bias but not sufficient to overpower the negative bias on the shield grid. The a-c firing peaks are generated in the peaking transformer having its primary in the phase-shift circuit on the left. The peaks are timed by the heat-control resistor $R10$. The weld is made by opening the weld-initiating switch $S1$ which action removes the d-c negative bias on the shield grid and permits tube 2 to conduct on the next voltage peak in transformer $T2$. Capacitor $C2$ discharges through the ignitor of tube 3 and a single half-cycle pulse from the power supply passes the welding transformer. When $C2$ discharges the potential on the anode of tube 2, it drops to a low value so that the tube ceases to conduct. Simultaneously the voltage across $C3$ rises to a high value which also prevents conduction. When the initiating switch is released (closed) the negative d-c bias across the shield grid is restored to prevent conduction of tube 4. Following the release of $S1$, the potential across $R2$ and $R3$ is restored to normal and $C2$ becomes charged and ready for a repeat cycle of operation. The voltage-regulator tube 4 is used to limit the transient negative bias placed across the shield grid when the initiating switch is released. A typical half-cycle bench welder with heat control is illustrated in Fig. 25.

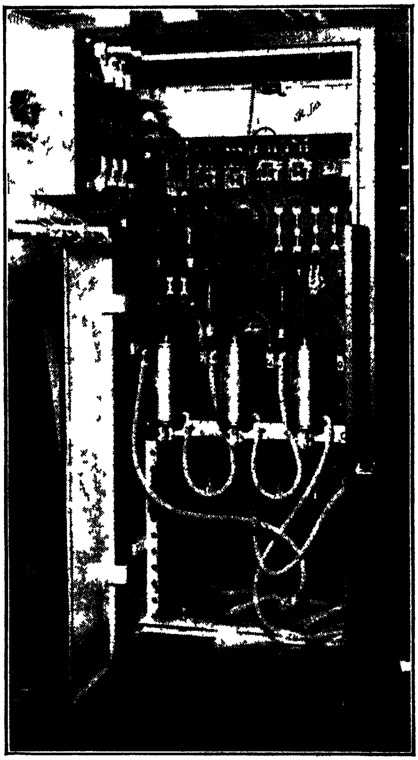


FIG. 27. Ignitron power rectifier for use with magnetic energy-storage spot welders. (Courtesy General Electric Company)

A second type of spot-welding control circuit for bench welders is shown in Fig. 32. This control uses thyratrons for closing the power circuit and is adjustable for timing spot welds of several cycles' dura-

tion. The explanation of the operation of this circuit constitutes one of the problems at the end of this chapter.

Energy-Storage Welding Systems. Two methods of energy storage for the welding of nonferrous metals have been developed. These are known as magnetic storage and electrostatic storage. The specific advantage of these methods is the reduction in the heavy transient peak demands on the power supply system and the accompanying voltage fluctuations. Large peak demands are created when nonferrous metals are welded because, when heated, these metals pass quickly from the solid to fluid state. During this transition period

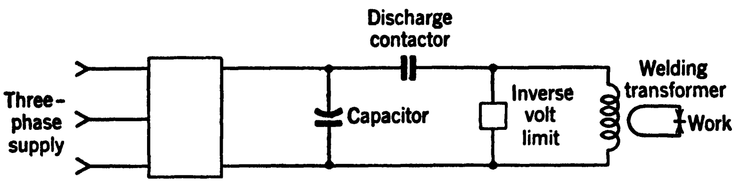


FIG. 28. Diagram for an electrostatic energy-storage spot welder.

there is only a narrow temperature range within which they exhibit a plastic tendency and can be welded successfully. Accordingly the welding process requires a high rate of concentrated heating which must be carefully controlled. These statements apply especially to the welding of aluminum and its alloys which have been used extensively in airplane construction.

The *magnetic energy-storage system* of welding utilizes the ability of a magnetic field to store electrical energy. The energy is stored in the magnetic field of the welding transformer which differs from the conventional welding transformer in having a core with a large cross section and an air gap in the circuit. The primary of this transformer is energized by a d-c voltage (Fig. 26*a*) resulting in a current having the exponential rising characteristics of an inductance excited by a d-c voltage. This rise of primary current is shown in part *b* of the figure. Since the rate of change of primary current and the resulting flux is relatively slow, the corresponding secondary current is small and produces little heat in the work. When the primary current reaches the proper value to give the desired storage energy (energy = $\frac{1}{2}LI^2$) the primary circuit is interrupted. After the primary circuit is opened the energy stored in the magnetic field produces a undirectional peaked pulse of discharge current in the transformer secondary winding and welding load. The integral (I^2RT) of this pulse of current heats the

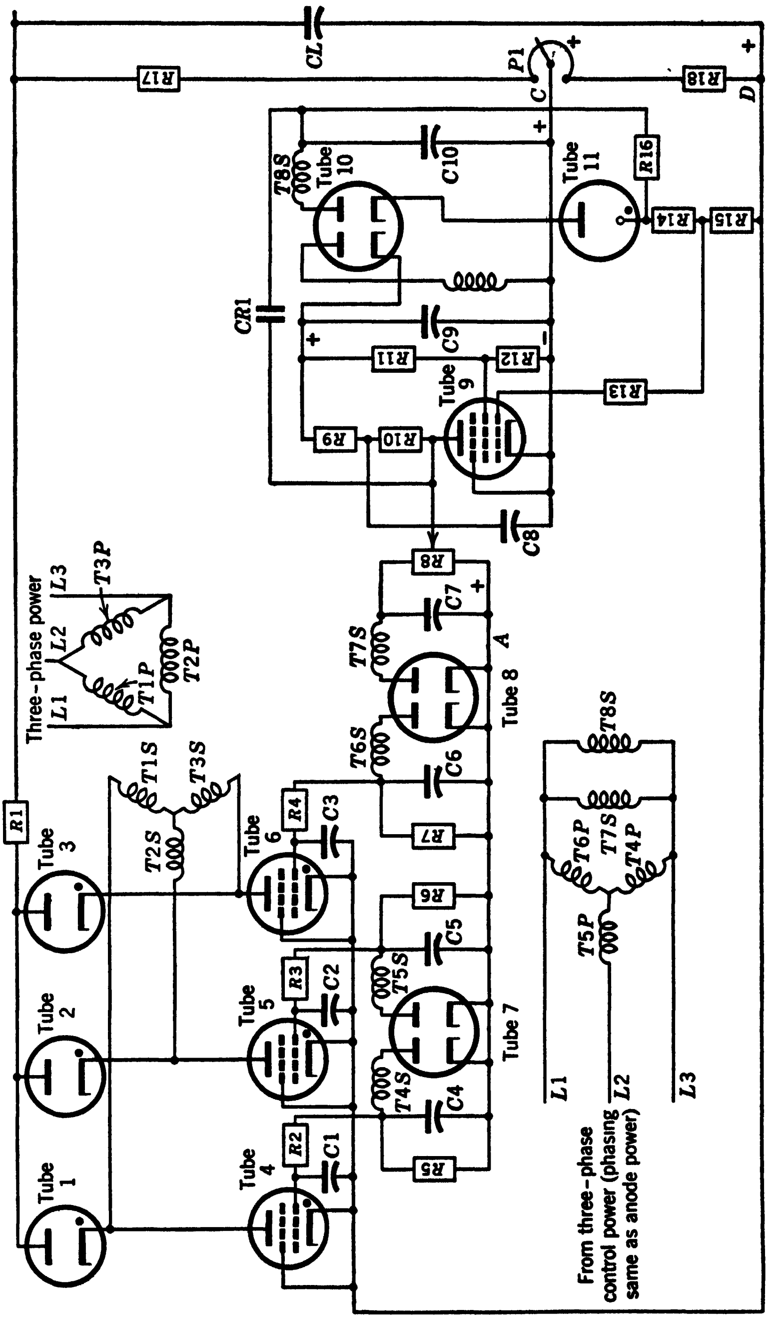


FIG. 29. Rectifier and voltage-control circuit for an electrostatic energy-storage spot welder.

work and produces the weld. The primary circuit of the weld transformer is not interrupted by a simple switch as shown in Fig. 26*a* but by a series of contactors so arranged that they may be progressively and automatically operated to introduce series resistors until operation of the last contactor completely opens the circuit. The contactors are actuated by a series current relay which may be adjusted over a wide range to govern the energy stored for the weld.

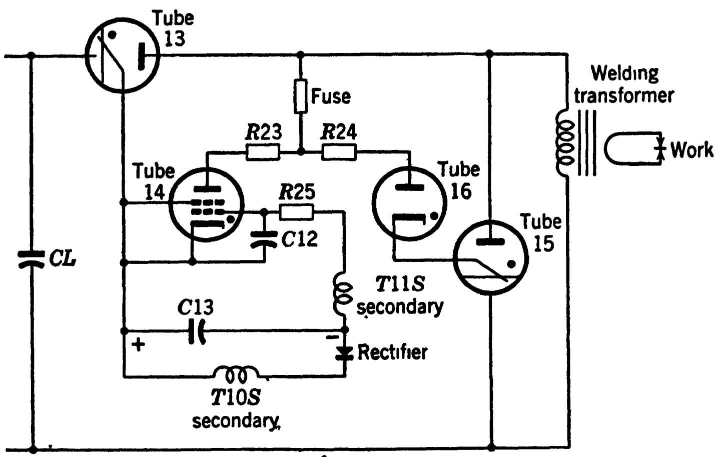


FIG. 30. Discharge oscillation suppressor circuits of an electrostatic energy-storage spot welder.

The d-c supply for the magnetic energy-storage system utilizes potentials varying from 130 to 170 volts and peak currents of 200 to 1500 amperes. Since few industrial plants have a d-c supply of this capacity available it is customary to furnish the d-c supply by a three-phase rectifier using one ignitron fired by one phanotron in each phase. For operations of large welders of the magnetic energy-storage type, an adjustable rectifier output becomes desirable. Such control may be attained by firing the ignitrons in the three-phase rectifier by individual thyratrons having phase-shift grid control. A rectifier unit of this type is illustrated in Fig. 27. In practice a single rectifier unit is utilized to operate two welding machines having their circuits interlocked to prevent simultaneous welding and to smooth out the load on the power supply system.

The *electrostatic energy-storage system* makes use of the ability of a capacitor to store electrical energy ($\frac{1}{2}CE^2$). Such a system requires that a capacitor bank be charged to a predetermined voltage and then

discharged through the metal to be welded. The total capacitance, the voltage to which it is charged, and the inductance and resistance of the discharge path determine the amount of available energy and the

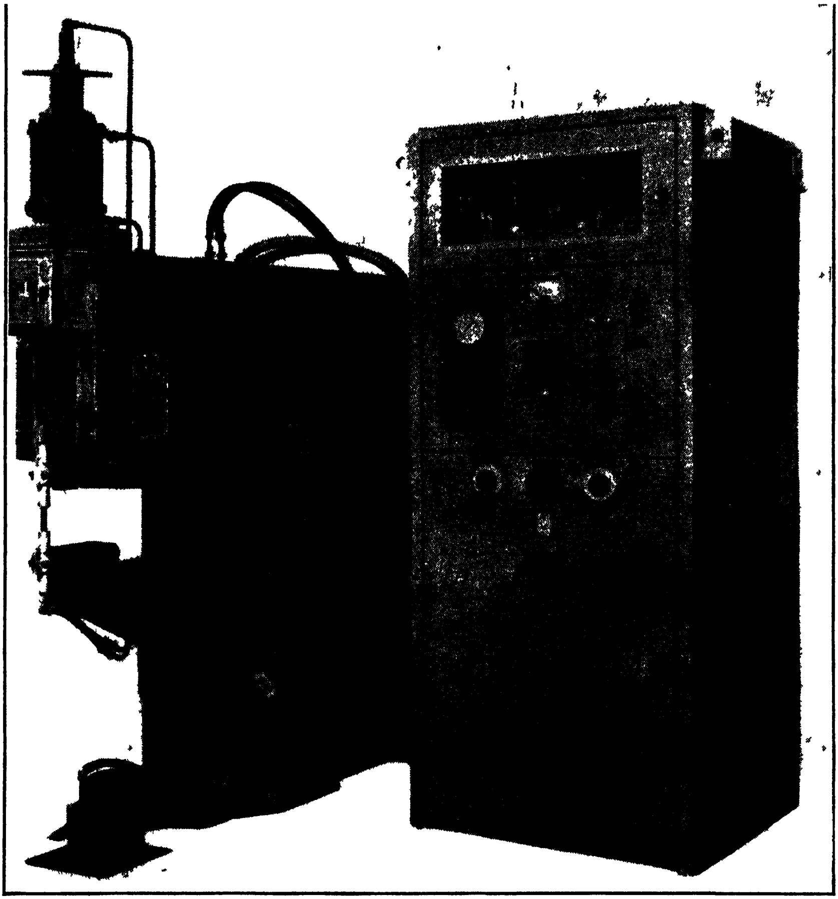


FIG. 31. Capacitor discharge control applied to a stored-energy-type welder which handles up to two thicknesses of 0 081-inch aluminum. (Courtesy General Electric Company)

rate at which it is delivered to the weld. Since capacitors are costly and large in size for low voltages, it is customary to place the capacitors in the primary circuit of the welding transformer and to use rather high d-c voltages since the stored energy varies as to the square of the voltage. A block diagram for an electrode energy-storage system is given in Fig. 28. The value of capacitance used in the primary circuit

ranges from 120 millifarads upward and the charging potentials applied to the plates vary from 1000 to 3000 volts direct current. The rectifier and voltage-control circuit for charging the discharge capacitor differs from those in preceding discussions because the voltage involved is

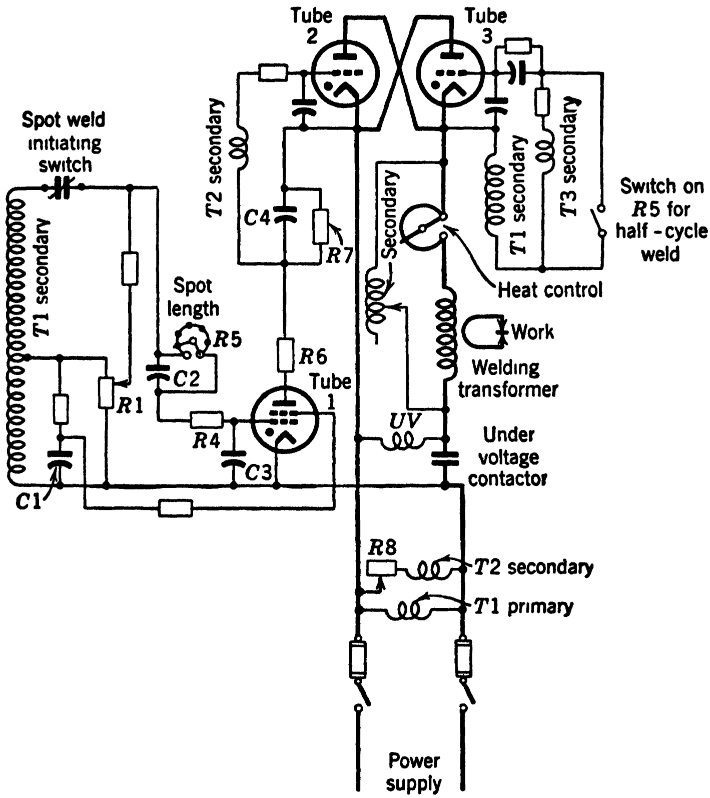


Fig. 32

higher and because the storage capacitor acts as a virtual short circuit. A schematic diagram for this special circuit is given in Fig. 29. The transformers for supplying the three-phase rectifier deliver 3800 volts, line to line, and are designed with a sufficient reactance to limit the short-circuit current to approximately 15 amperes direct-current. The rectifying tubes are phanotrons which are in series with thyratrons for controlling the rate of charge and the ultimate voltage to which the storage capacitor is charged. The circuit of Fig. 29 permits the thyratrons to conduct over a possible angle of 240 degrees instead of the 180 degrees found in the half-wave rectifier.

The discharge circuit for the storage capacitor of the electrostatic system is given in Fig. 30 which is a continuation of Fig. 29. This circuit is designed to prevent any reversal of the capacitor voltage in case an *R, L,* and *C* combination might exist which would tend to cause oscillations. If the reverse voltage reaches a value of 100 to 150 volts, ignitron tube 15 is made to conduct. The explanation of the operation of this circuit is left to the reader.

In addition to the electronic circuits shown, the complete electrostatic storage system includes sequencing controls, protective equipment for overload, over-voltage, and reversing contactors which reverse the connections of the welding transformer on alternate welds to prevent accumulative effects of discharge and possible saturation of the transformer. An electrostatic energy-storage welder and controller are shown in Fig. 31.

Problems

1. In Fig. 16, capacitor $C1$ and resistor $R5$ are set to give a time delay of 30 cycles in dropping the voltage across the capacitor to 37 per cent. If $C1$ is 1 microfarad, what is the value of $R5$? Solve the problem for 5 cycles of timing.
2. Explain the theory of action in the oscillation-suppression circuit of Fig. 30. (Thyratron tube 14 is triggered by an impulse voltage into $T11S$).
3. Explain the action taking place in the welding circuit shown in Fig. 32.

REFERENCES

Gillette, R. T., "Resistance-Welding Methods and Equipment," *General Electric Bulletin GET*-1189.

Palmer, H. L., M. E. Bivens, S. A. Clark, G. L. Rogers, and Barton L. Weller, "Electronic Welding Control," *General Electric Bulletin GET*-1170.

Electronic Laboratory Manual, Westinghouse Electric Corporation.

Chapter XV

ELECTROSTATIC PRECIPITATION

Electrostatic Precipitation. Particles of dust, smoke, condensed vapors, and chemical fumes may be removed from gases by mechanical processes and by electrostatic precipitation. The mechanical processes utilize centrifugal devices, simple filters, washers, and viscous-coated devices which depend on impinging the particles upon a coated surface. Electrostatic precipitation is the process of removing suspended particles from gases by the aid of an electrical discharge between electrodes in the gas stream. The cleaning of gases by a corona discharge was suggested by Hohlfield in 1824. In 1906, Dr. Frederick G. Cottrell, a professor in physical chemistry at the University of California, developed a practical electrostatic system which has been very successful in handling difficult cleaning problems that could not be accomplished by other methods.

The fundamental principle of electrostatic precipitation lies in the ionization of the gas by a strong electrical field accompanied by a corona discharge. The ionized gas particles consist of electrons, negatively charged ions, and positively charged ions. The electrons and negatively charged ions become attached to suspended particles in the gas, giving them a negative charge. In the usual form of precipitator, the negative electrode is of comparatively small diameter to give a concentrated field of high gradient, and it is placed at a suitable distance from a large positive electrode which serves as the collector. The negatively charged particles are attracted to the positively charged electrode and collect on its surface. The particles thus collected either drop by gravity into a collecting hopper below, or after interrupting the electric supply the electrodes are rapped manually or automatically to cause the precipitated material to be dislodged and fall into a hopper.

Cottrell Electrical Precipitator. The equipment used in the Cottrell electric precipitation process is shown schematically in Fig. 1. The precipitating chamber at the left is supplied with high-voltage direct

current through the action of a rotary converter which rectifies the alternating voltage supplied by the transformer. The simplest precipitator unit (pipe type) consists of an insulated wire suspended at the center of a tube or pipe. A high negative potential is supplied to the wire and the grounded positive pipe constitutes the other terminal for the system. Large-capacity precipitators frequently consist of a large number of tubes and wire electrode units placed in parallel. A second general form for the Cottrell precipitators is illustrated in Fig. 2.

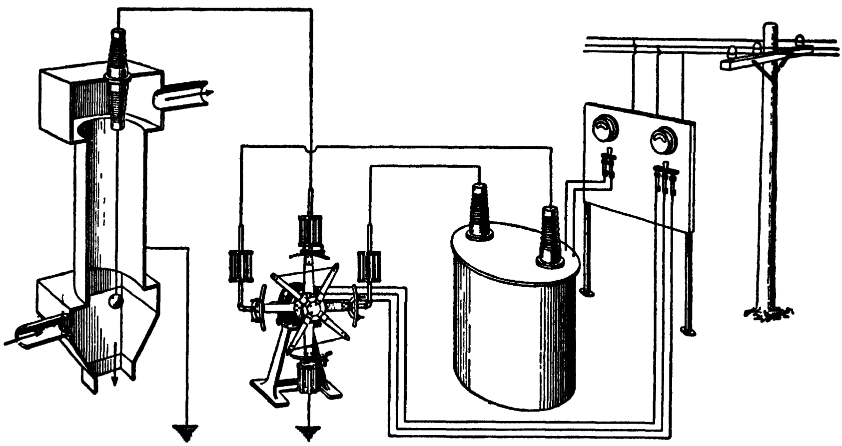


FIG. 1. Schematic diagram of the equipment used in the Cottrell electric precipitator. From left to right: precipitator, rotary rectifier, transformer, switchboard, and supply.

Here a number of wires are suspended between parallel plates which constitute the grounded collecting electrodes. Several designs for the plate collectors are in use and are constructed as follows:

1. *Solid steel,* sometimes corrugated.
2. *Concrete,* also called graded resistance, with conductors imbedded in the center.
3. *Rod curtain,* in which curtains of small rods or pipes are hung close together to form the collecting electrodes.
4. *Perforated plate,* where the collecting electrodes are in the form of narrow boxes with perforated sides.
5. *Pocket type,* in which the precipitated material is trapped in pockets, usually with upward gas flow.

The cross section of a heavy-duty precipitator using parallel plates and vertical flow is shown in Fig. 3.

In order to produce the corona discharge of sufficient magnitude for the Cottrell precipitation process, potentials of the order of 30,000 to 75,000 volts are required. These high potentials are obtained by rectifying the output of a high-voltage transformer. Rectification may be accomplished by special high-voltage kenotrons (see Fig. 16, page 53)

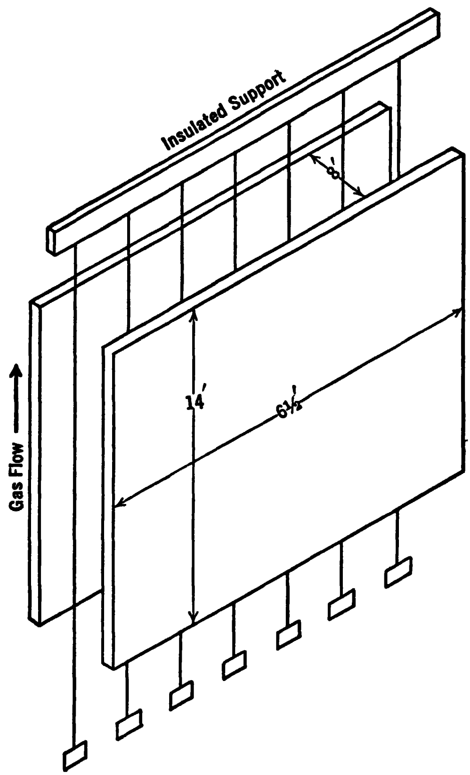


FIG. 2. Electrode assembly for electrostatic precipitation.

employed in a single-phase or three-phase rectifier circuit or by commutating rectification. Nearly all large-scale installations have employed rotary rectifiers because of a lower initial cost and a lower cost of maintenance. Rotary rectifiers are illustrated in Figs. 1 and 4. A three-phase, four-pole (induced-pole type) synchronous motor carries a Bakelite rotor disk on which are mounted four rotating contact tips connected in pairs by quadrant conductor strips imbedded in the Bakelite. The rotating contact tips pass under stationary shoes connected to the transformer and precipitator unit in such a way as to re-

verse the connections from the transformer to the precipitator in synchronism with the a-c wave and thus produce rectification. Provision must be made for testing the polarity of the rectified wave and for slipping poles to assure correct polarity when starting the motor.

Some arcing occurs at the rotary rectifier and within the precipitator which will cause some radio interference unless corrective measures are taken. Radio interference or corrective coils are usually placed directly at the terminals of the shoes on the rotary converter as shown in Fig. 4.

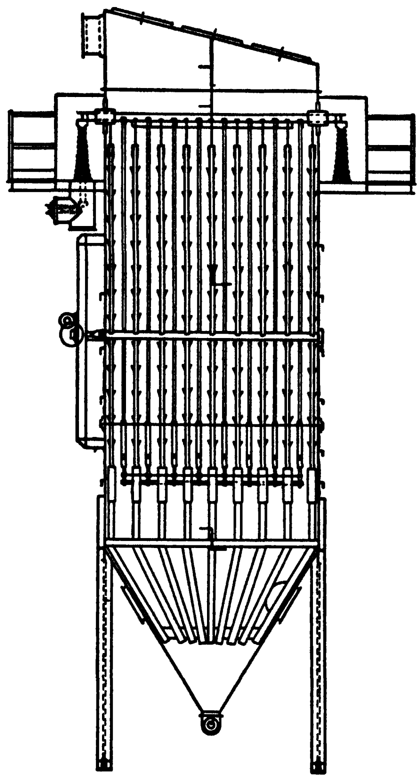


FIG. 3. Vertical-flow high-duty-type precipitator. (Courtesy Western Precipitation Corporation.)

Transformers for the Cottrell process are of a special surge-proof design. They are furnished in ratings from 7.5 to 25 kilovolt-amperes, depending on the requirements of the load. Taps are provided on the secondary so that the voltage may be held at the maximum value to suit the particular operating conditions. The voltage delivered to the precipitator is limited by the severity of the sparking in the precipitator, which, in turn, depends on the density and nature of the dispersoid, the temperature, and the water-vapor content.

Applications of the Cottrell Process. The Cottrell precipitators have been applied in a large number of manufacturing processes. Blast furnaces give forth enormous volumes of gases containing hundreds of tons of dust. Much of the exhaust gas is burned later in stoves, open-hearth furnaces, and gas engines where impurities in the gas may be damaging. Cottrell precipitators are used for preliminary and sometimes for final cleaning of the gas. In smelter operation precipitators are used for recovering from gases such valuable materials as gold,

silver, cadmium, lead, and zinc. The value of such recoveries has amounted to $40,000 to $1,000,000 gross value per year, showing the economic importance of the process. Precipitators are used to remove the fog and water mist from manufactured gas after it has been scrubbed and cooled. The exhaust gas from cement plants gives a heavy dispersion of dust over the surrounding area. Electric precipitators will remove this dust, making it possible to return the recovered material to the kilns or to separate the potash from the recovery and sell it for fertilizer. In oil refineries precipitators may be used to collect sulphuric acid mist that is carried over in gases. In catalyst-cracking plants for producing high-octane gasoline a costly catalyst may be recovered by the precipitation process to the extent of 99 6 per cent Power-generating plants using pulverized coal for fuel disperse large quantities of fly ash over the surrounding territory. Electric precipitators are collecting from 90 to 98 per cent of the total fly ash in more than a hundred powdered-fuel plants in the United States and in many hundreds in other parts of the world. A fly-ash precipitator installation on a power plant in Argentina is shown in Fig. 5.

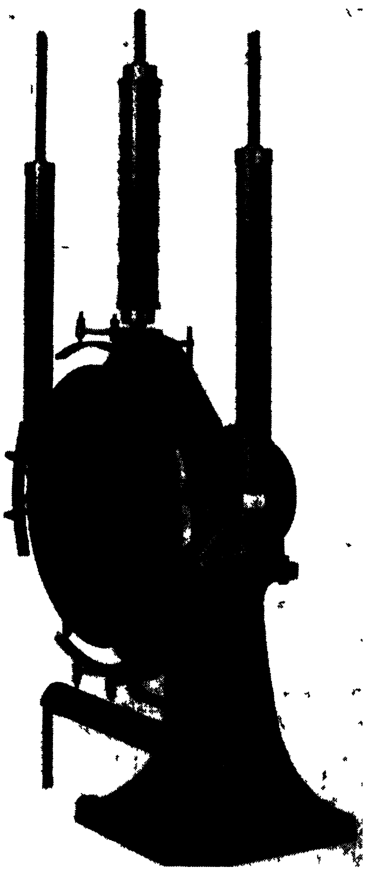


FIG. 4. Rotary rectifier with radio interference suppressor coils for electrostatic precipitation. (Courtesy Western Precipitation Corporation.)

Some of the newer applications of electrostatic precipitators are (1) the application of sand and Carborundum to the backing material in the manufacture of abrasive cloth and (2) the placement of phosphor material in a cathode-ray tube.

Electrostatic Precipitation for Air Conditioning. A new and valuable method for applying the principle of electrostatic precipitation has been perfected through the research and developmental work of

G. W. Penney. This new method makes possible a satisfactory and economical process for purifying air for human consumption as a part of a complete air-conditioning system. The process of electrostatic precipitation described in the preceding discussion is not satisfactory

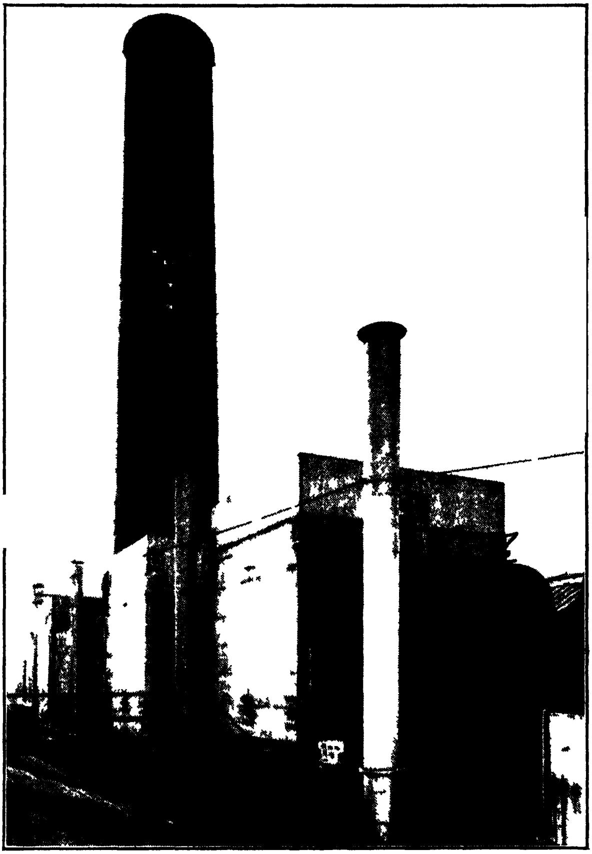


FIG. 5. Fly-ash precipitator at an Argentine power plant. (Courtesy Western Precipitation Corporation.)

for air conditioning because of the high voltages of 30,000 to 75,000 volts used. Such high voltages result in a high cost of operation because of (1) relatively large power consumption, (2) high initial investment for equipment, and (3) large maintenance expense. Of still greater importance is the fact that the method described produces so much ozone that the cleaned air, although free of dust, is too irritating to the nose and throat to be used where human beings are concerned.

In the new method of precipitation, the functions of charging the particles by corona discharge and the collection of the charged dust particles are performed in two separate stages of equipment. This new method utilizes a lower voltage for the ionization and charging process and a still lower voltage for the collection of charged dust particles.

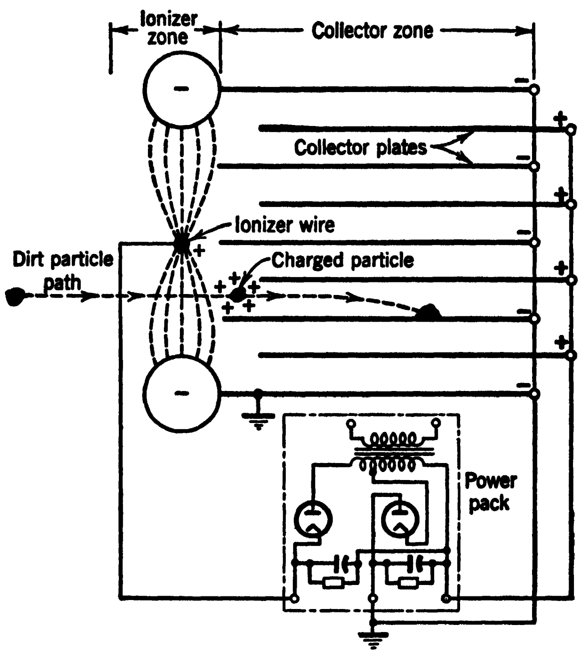


FIG. 6. Construction and action in a Precipitron. (Courtesy Westinghouse Electric Corporation.)

One of the principal advantages of separating the functions is the large reduction in the ionizing current which causes a reduction in ozone generation. The construction for separating these functions is shown in Fig. 6. In the ionizer zone at the left a fine wire electrode carries a d-c positive potential of 13,000 volts and is located between grounded cylinders (electrodes). In this zone the gas is ionized and the dust particles are charged while they are being swept into the collector zone. The collector zone consists of many plates disposed edgewise to the air flow. Alternate plates are charged positive to about 6000 volts, with the intervening ones held at zero or ground potential. The positively charged dust particles are collected on the negative or grounded plates where they are held until removal.

The theory of the action taking place is of interest. In the ionizer zone a very high potential gradient exists close to the positively charged ionizer electrode. This potential is sufficient to extract electrons from the surrounding atoms of gas. Having lost electrons the atoms become positive ions and are moved by the existing potential gradient along the lines of the electrostatic field toward the negative electrodes in the ionizer zone. Since dust particles usually have a dielectric constant greater than one, the lines of the electrostatic field will be distorted with more lines passing through the particle. This distortion causes more of the positive ions to collide with and become attached to the dust particles, thus giving them a positive charge very quickly. The charging of the dust particles is also aided by the kinetic or heat motion (random) of the positive ions which results in more collisions and hence in faster charging of the particles. The dust particle will continue to acquire a positive charge from the + ions until the magnitude of its charge exerts a repelling force sufficient to prevent further attachments of charge.

After the dust particles are charged, a few of them along with some unattached + ions may reach the negative electrodes in the ionizer zone. However, most of the dust particles will be carried by the gas stream into the collector zone. In this collector zone there exists between the plates a high potential gradient arising from the 6000 volts applied to alternate plates. This potential gradient E acts upon the charged particles with a force equal to E times the charge Q on the particle. Thus the particle is urged toward the negative plate but is opposed by the air resistance with the resulting motion shown in Fig. 6. Obviously, the magnitude of these factors and the length of the collector plates will determine the maximum permissible velocity of the air stream which will assure that the charged particle reaches the collector plate before emerging from the collector zone. It is well to note that this dual-stage system provides (1) a nonuniform field for ionization which is most efficient for this purpose and (2) a high-magnitude uniform field for collection of charged particles which, in turn, is most rapid and efficient for its function.

When the charged particle reaches the collector plate it gives up most if not all its charge if it has conducting properties. Hence electrostatic forces cannot be relied upon for retaining the particles. In general, the molecular force that exists between substances in contact offers a form of adhesion which will hold the particles to the collector plate. In special cases the plates may be coated with oil or other material which will cause particles to adhere, although they may ad-

here so well that cleaning of the plates becomes rather difficult. With some types of dust, the particles may adhere normally but a jar may loosen agglomerations of dust. This problem may be solved by using a simple mechanical filter beyond the precipitator. The method used

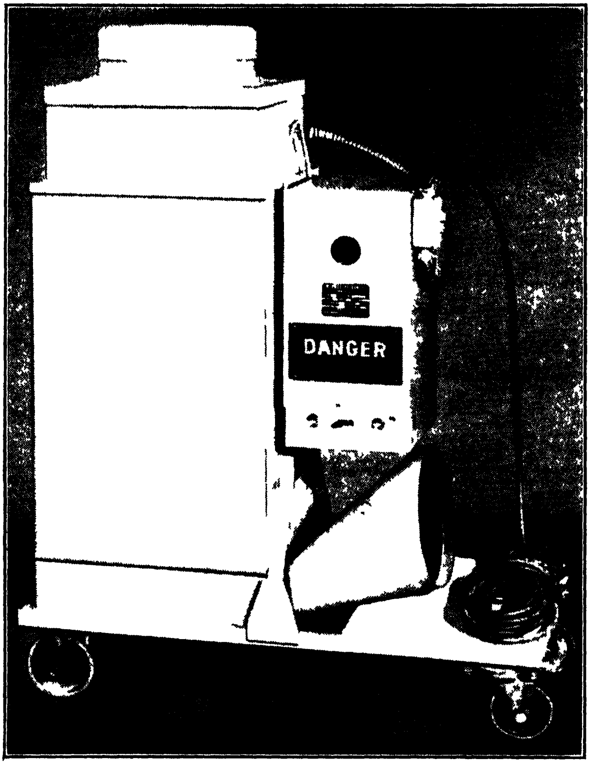


FIG. 7. The Precipitron, a portable precipitation unit. (Courtesy Westinghouse Electric Corporation.)

for cleaning the plates depends on the nature of the dust collected. Cleaning is usually accomplished by a direct washing process.

The Precipitron. The principle of electrostatic precipitation described is used in commercial assemblies known as the Precipitron. These assemblies are made up of standard sections combined to give the desired air-cleaning capacity. Individual Precipitron units range in capacity from 240 to 350,000 cubic feet per minute. A small portable unit is shown in Fig. 7. A simplified circuit for the rectifier which supplies the high voltage for the Precipitron is given in Fig. 8. The transformer steps up the line voltage to 7000 volts and delivers it

through a voltage doubling arrangement (see page 278) to two kenotron tubes which rectify it for use in the Precipitron. A part of the necessary indicating and protective equipment for the rectifier is shown at the left of the circuit diagram.

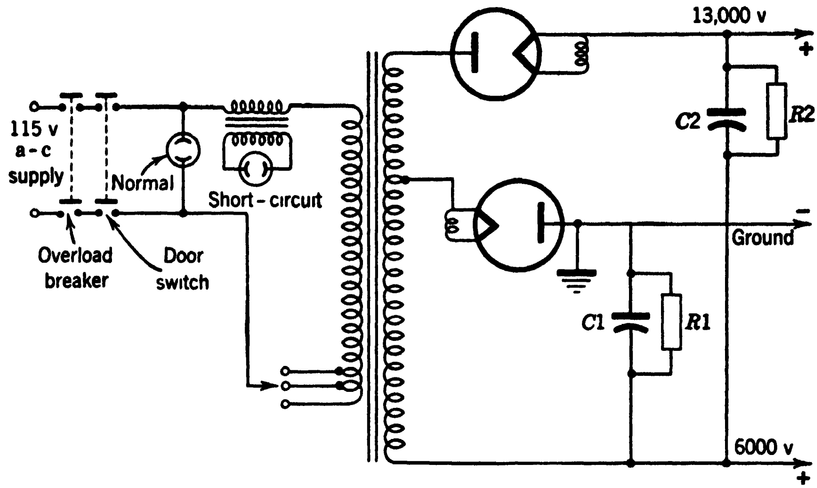


FIG. 8. Basic circuit of a Precipitron.

The Precipitron is particularly adapted to removing light concentrations of fine dust. Some of its principal fields of application are:

1. Removal of industrial dusts which constitute a hazard to the health of employees.
2. Air cleaning to protect delicate apparatus or processes.
3. Air cleaning in homes and offices in soft-coal-burning cities to reduce cleaning of walls and draperies.
4. Air cleaning for the relief of hay fever and asthma.
5. Air cleaning in stores to reduce damage to merchandise.

REFERENCES

Cottrell Process of Electrical Precipitation, Research Corporation, and Western Precipitation Corporation, 1940.

SCHMIDT, WALTER A., and EVALD ANDERSON, "Electrical Precipitation," *Elec. Eng.,* August 1938.

PENNEY, G. W., "A New Electrostatic Precipitator," *Elec. Eng.,* January 1937.

Chapter XVI

ELECTRONIC OPERATION OF DIRECT-CURRENT MOTORS

The a-c system presents well-known advantages for the generation, transmission, and distribution of electric energy which has caused its universal adoption throughout the world. Most a-c motors are rugged and simple in construction, cost relatively little, and require little maintenance. These motors are inherently of the constant-speed type; they have low starting torque, high starting current, and a low power factor. Direct-current motors are more costly and require more maintenance but they do possess excellent characteristics of speed and torque. The d-c shunt motor can be operated over a wide range of speed with good regulation at all speeds and it possesses good torque characteristics. The d-c series motor develops high starting torque with a variable speed and a nearly constant energy input, which are desirable characteristics for certain types of loads. These desirable characteristics of d-c motors have caused an increase in the use of d-c machinery in industry since 1935 with the result that during World War II approximately one-half the kilowatt capacity of all machinery manufactured above one horsepower was direct current. The use of d-c motors on the a-c distribution system requires the use of some form of a-c to d-c conversion equipment.

The systems of a-c to d-c conversion equipment in general use for d-c motor drives are:

1. D-c constant potential by motor-generator sets.
2. D-c variable or adjustable voltage by motor-generator sets.
3. Electronic conversion methods.

System (1) is used in industries where electric cranes and hoists using series motors and perhaps shunt or compound motors are required. System 2 is used for electric drives for elevators, steel-rolling mills, textile mills, paper mills, and newsprinting machines. Since 1940 electronic conversion units for operating d-c shunt motors giving

a wide range of speed control with excellent regulation and automatic features have been available. These units have been built for capacities of 1 to 200 hp and it is feasible to design larger sizes when needed.

A brief review of the factors controlling the starting and speed control of d-c shunt motors will simplify the theory of the electronic units for operating the machines from an a-c supply. A schematic circuit and the speed characteristics of a shunt-motor arc are shown in

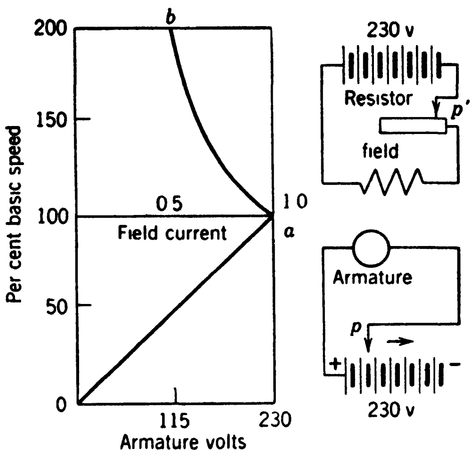


FIG. 1. Speed control of a shunt motor by varying the armature volts and field current.

Fig. 1. For starting, a full field (field volts and current) and a reduced voltage across the armature is required to give maximum torque and a limited current through the armature. The armature current may be limited (for zero or low counter emf) by a series resistor on constant-potential systems or by application of a low voltage. This can be accomplished by the setting of p and p' in the schematic circuit of Fig. 1. As p is advanced to the right the voltage impressed across the armature will rise and the speed will increase linearly from 0 to a as the impressed armature voltage rises to its normal value of 230. This will give the basic or 100 per cent speed. Now the speed can be raised further by reducing the field flux and current through a movement of the point p' to the left along the field resistor. The rise in speed is inversely proportional to the field flux and will follow along the line ab. By this sequence of adjustments any speed from zero up to the upper limit of motor design may be attained and the speed regulation for any given setting will be good.

A simple schematic circuit for applying an electronic converter to a shunt motor is given in Fig. 2. Here a unidirectional current is supplied from an a-c source by two rectifier diodes. The motor is started manually by a starter (resistor) in the armature circuit to protect the armature and the diodes. After basic speed is attained the speed may be raised higher by manual operation of a field resistor as on any constant-potential supply system. The obvious objection to this circuit and system is that the protection and life of the equipment is dependent upon the judgment and care of the operator.

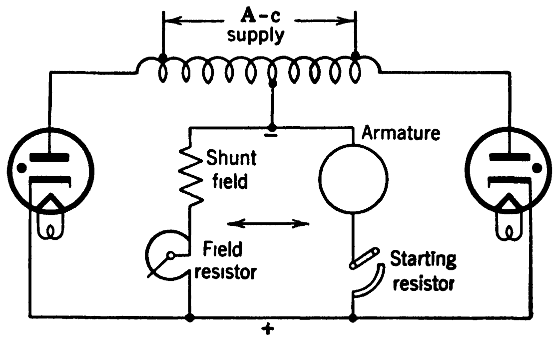


FIG. 2

The simple circuit of Fig. 2 can be modified by the addition of some protective units and made automatic as shown in Fig. 3. The unit is put in operation by closing the double-pole switch on the a-c supply. The closing of this switch energizes the filament transformer and starts the warm-up of the rectifier tubes. Simultaneously, a time-delay mechanism and relay (motor *M* and *TD*) begins the measurement of the time necessary for the preheat of the tubes. When this time has passed the time-delay relay operates its contacts which energize the anode-contactor relay *ANC* and connect the rectifier to the motor. The thyrite resistors across the shunt field and the secondary of the transformer snub any transient inductive voltage kick that may occur when the main switch is opened. Since the motor must start with full voltage applied to the armature, this circuit is applicable only to fractional horsepower motors which have an armature resistance and inductance high enough to permit such service. A commercial form of electronic controller using this circuit is illustrated in Fig. 4. This type of electronic unit permits a speed adjustment of the order of three to one above the basic speed of the motor through the operation of shunt field resistor *F*.

For motors of 1 to 5 hp where an adjustable speed control by field above basic speed is satisfactory, the circuit of Fig. 3 needs to be modi-

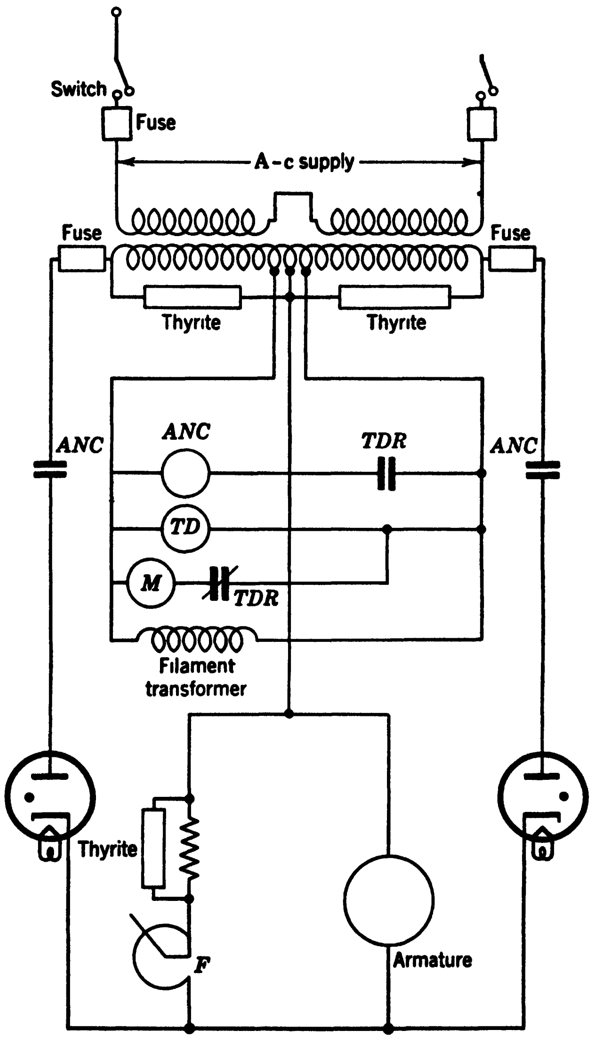


FIG. 3. Simple electronic rectifier circuit for operating a d-c motor.

fied by the addition of a magnetic time starter with push-button control. In this circuit a starting resistor is inserted in series with the armature and a second resistor for dynamic braking is employed. The timing for the preheat of the rectifier tubes is interlocked with the

starter timing to prevent starter action before the tubes are conditioned for motor starting.

Many motor applications require a wide range of speed control from zero to a maximum value. This speed range necessitates a variation of armature voltage for speeds from zero to basic speed. This necessity can be met by substituting the thyratron for the diode rectifier and applying the proper phase shift to the grid voltage. A simple schematic circuit illustrating this application is given in Fig. 5. Here the armature is fed by thyratrons *a* and *b* and the shunt field by thyratrons *c* and *d*. The phase shifters 1 and 2 and similarly 3 and 4 are mechanically locked together. Thus, to operate the motor after the necessary preheat of the tubes, phase shift 3–4 should be set for full field current and the phase shift of 1–2 should be 180 degrees, so that tubes *a* and *b* are inoperative. Then phase shifts 1–2 can be advanced to raise the voltage across the armature and bring up the speed. For speeds less than basic armature phase shift is applied, whereas for speeds above basic the field phase shift is adjusted after the armature is adjusted to full normal voltage. This control of speed can be visualized by reference to the speed characteristic curve in Fig. 1. The simple circuit of Fig. 5 illustrates the basic principle of electronic motor drives. However, since it assumes manual control, other and more complicated circuits are employed to give the necessary protective, automatic, and "electronic brain" features of control.

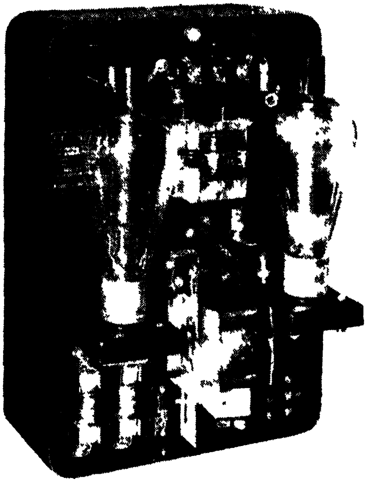


FIG. 4. Thyratron motor controller for fractional-horsepower d-c shunt motor. (Courtesy General Electric Company.)

The complete electronic operation of a d-c motor will embody the features of wide range of speed adjustment, constant speed for any adjustment, current limit for armature current, automatic starting, motor reversing, motor jogging, and dynamic or regenerative braking. Some applications will not require all these features but the complete control unit will require the use of the majority of the circuit and equipment components previously discussed in Chapter X. The exact

circuit, equipment, and arrangement used may vary with the equipment manufacturer. Hence the following discussion should be considered as typical for this form of electronic control.

The speed of a d-c shunt motor is governed by the voltage impressed across the armature. This voltage may be supplied and controlled by

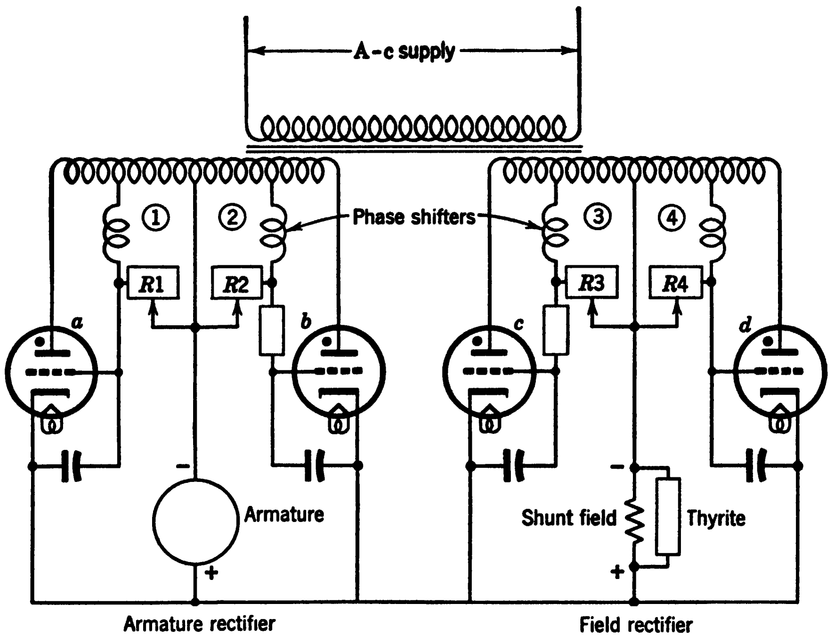


FIG. 5. Schematic electronic starting and speed control for a shunt motor using thyratrons with phase-shifting control.

the simple circuit of Fig. 6. Here the phase shift is produced by the bridge circuit having as one arm a saturable reactor *SR*. (See pages 222 and 225.) The complete control from "full off" to "full on" is accomplished in a stepless manner and with a small current in the d-c winding which is supplied by a triode amplifier. Before discussing the amplifier circuit another component circuit should be recalled. Since the applied a-c voltage may vary with the motor load and with other load and line conditions, it is necessary to provide a constant or reference source of voltage if speed and other desired characteristics are to be maintained constant. Such a constant source may be provided through the use of the constant-voltage circuit explained on page 234. The three-wire constant-voltage source of Fig. 7 now becomes a basic

part of the circuits which follow. An amplifier triode *D* is connected to the constant-voltage system and to the saturable reactor *SR*, as shown in Fig. 8 (right). The cathode-to-plate plus *SR* circuit is across 75 volts (constant) and the grid of triode *D* is tapped to a potential divider circuit between *R*3 and *R*4 so that normally it is sufficiently positive so that tube *D* conducts fully and *SR* advances the

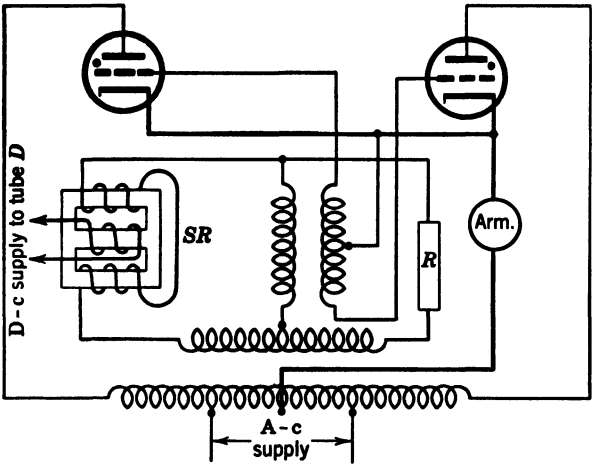


FIG. 6. Phase-shifter circuit applied to motor armature voltage supply.

thyratrons to "full on." In order to control tube *D* it will be necessary to vary the potential of its grid in some manner. Such control is accomplished by a second triode *C*. To understand the action of the remainder of the circuit it should be recalled that under stable operation the speed of the rotating armature is proportional to its counter emf and approximately proportional to the impressed emf. The armature voltage is impressed across the potential divider circuit consisting of resistance *R*9 and *R*10 (plus a small adjustable vernier resistance) having values such that the voltage drop e_A will be equal to 75 volts for the basic speed of the armature and full shunt field. The voltage e_A is to be compared with a standard (constant) voltage e_S obtained from a potentiometer resistance *S*. These two comparison voltages are applied to the grid and cathode of tube *C* which, in turn, controls the grid potential of tube *D*. To follow the action of the complete circuit, assume that the armature is at rest with e_A equal to zero and that a speed of ⅔ of basic is desired. Potentiometer *S* will be

advanced (CW) to raise e_S to 50 volts. Now, with the cathode of C at $+50$ volts and the grid at zero or -50 volts with respect to the cathode, tube C is biased below cutoff and does not conduct. Under these conditions the grid of tube D is positive, the tube conducts, and the SR bridge holds the thyratron rectifiers "full on." The rectified current flowing through the armature causes it to accelerate and, as it does so, its counter emf and e_A rise. As e_A approaches e_S in value, the grid of C becomes less negative and at a suitable value tube C begins to conduct electrons from cathode to plate. This electron current passes from the lower zero- or negative-voltage line through S, tube C, and resistor $R2$, to the 150 constant-voltage line. This new current through $R2$ increases the voltage drop over $R2$ and lowers the potential of the grid of tube D because the sum of the drops across $R2$, $R3$, and $R4$ are held constant at 150 volts. The lowering of the grid potential of triode D reduces the current output to the saturable reactor and shifts the phase of the thyratron grid voltages, thus reducing the rectified current and voltage to the armature. With proper adjustments of resistors and tubes, the armature speed is stabilized when e_A equals e_S. Should the mechanical load on the armature be reduced, the speed would tend to rise carrying e_A with it and increasing current output of tube C, which would make the grid of D more negative with subsequent retarding shift of phase of thyratron grid supply. It is obvious from this discussion that any motor speed from zero to basic can be attained by the setting of potentiometer resistor S. Thus S becomes the basic-speed adjuster control.

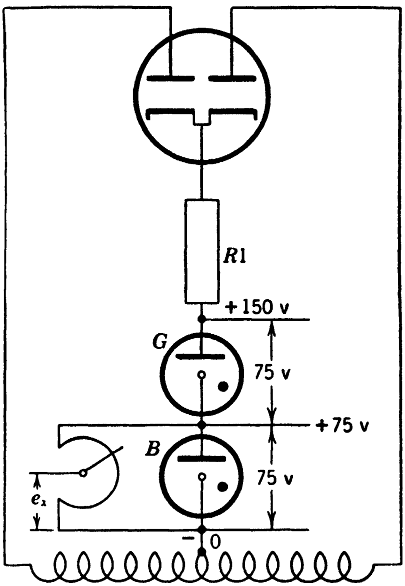


FIG. 7. Constant-voltage supply.

In the electronic control unit the excitation of the d-c shunt field will be obtained by a rectifier system. Speeds from basic up may be obtained by reducing the field flux and current by using a circuit identical with Fig. 8 except that the armature is replaced by the shunt field. In the new circuit the speed control (rise) will be secured by ad-

vancing S in the clockwise direction (instead of counterclockwise). Since the armature-voltage speed control is made with full field and since field control is accomplished with normal armature voltage, the control and equipment can be simplified by "ganging" or mounting both controls on one shaft and with one dial, as suggested in Fig. 9. Conductor segments are necessary on each control for the sectors in which the particular control is inactive.

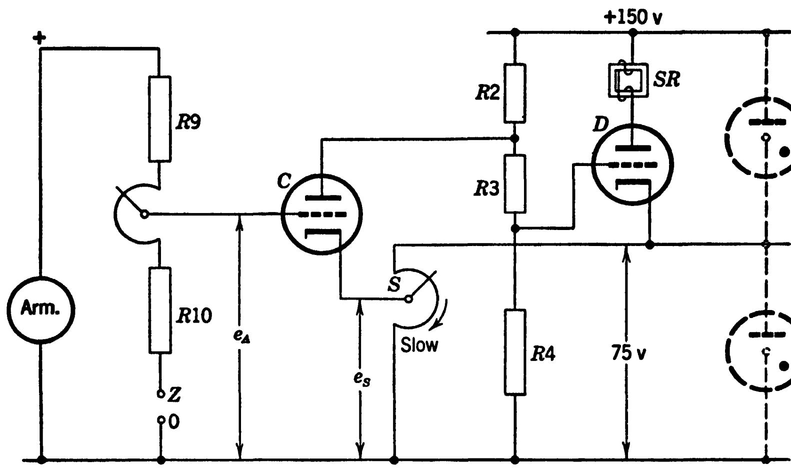


FIG. 8. Circuits for controlling the speed of an armature.

In the preceding discussion it was assumed that the voltage drop across the armature was approximately proportional to the speed and counter emf. Actually, the impressed voltage is equal to the counter emf plus the IR drop and the speed is proportionately lower than the impressed voltage. This means that the speed will be allowed to fall as the motor is loaded, the same as operation on a constant-voltage system. This drop in speed can be avoided if a voltage of the proper sign and equal to the IR drop in the armature were introduced into the circuit of Fig. 8 between points O and Z. The desired voltage can be secured from the rectifier circuit of Fig. 10. The primary P of the current transformer consists of two current coils connected into the anode circuits of the thyratrons which feed the armature. These coils are connected so that the rectified pulses of current flow in opposite directions and thus induce an alternating voltage across the secondary S. This secondary of the current transformer is loaded by a resistor R_Y to control its approximate voltage value. The transformer is also

loaded by a double-anode diode and an adjustable voltage drop is taken from the rectifier load resistance. The latter voltage drop which is proportional to thyratron load current is applied between points O and Z on Fig. 8. The rectifier circuit of Fig. 10 is also used to supply a voltage for current-limitation functions.

The preceding discussion and circuits show how a constant adjustable range of speed may be attained but omits the current-limitation function of electronic operation. A d-c motor armature during acceleration or during overload will take an excessive current unless protective measures are provided. An addition of one triode and circuit to the preceding circuit may be employed to give such protection. This current-limiter circuit is illustrated by tube E and its circuit in heavy lines on Fig. 11. The grid of tube E is fed from the same rectifier circuit that provided armature RI drop compensation. The circuit adjustment is such that the grid of tube E is biased negative for all armature-load currents up to, say, 150 per cent of normal load (other percentage values may be chosen). Then as the armature-load current rises above 150 per cent, the grid of tube E moves rapidly from slightly negative to positive and tube E conducts via resistor $R2$. The electron current through $R2$ lowers the grid voltage of tube D (like conduction of tube C) which, in turn, shifts the phase of the grids on the thyratrons. It should be noted that tube C functions through voltage comparison and at normal load values, whereas limiter tube E comes into action only for overload current values. Obviously, tube E functions equally well during armature acceleration and at all speeds and loads. The current-limiter action is also applied to the field circuit as a protective measure but in a reversed manner. Here

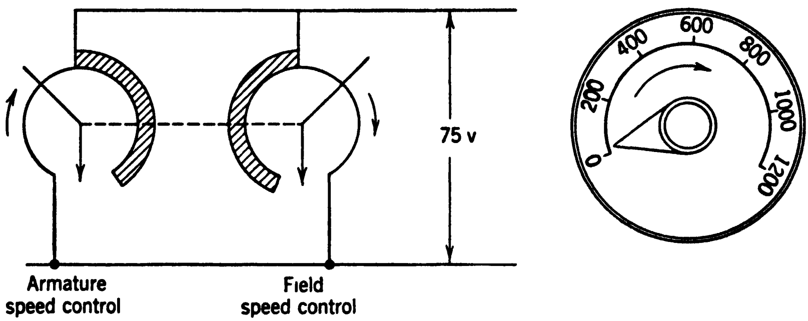


FIG. 9. Speed-control potentiometers mechanically locked together.

the grid of tube *EE* is connected to the same source as tube *E* but the rise of grid potential causes *E* to conduct and raise the potential of the grid of tube *DD* (increased drop across *R*25). The increased conduction of tube *DD* shifts the thyratron grids toward "full on." This action strengthens the field, raises the torque, and reduces the motor speed. It is customary to adjust the current-limiter circuits so that the field action precedes somewhat the armature current-limiter action,

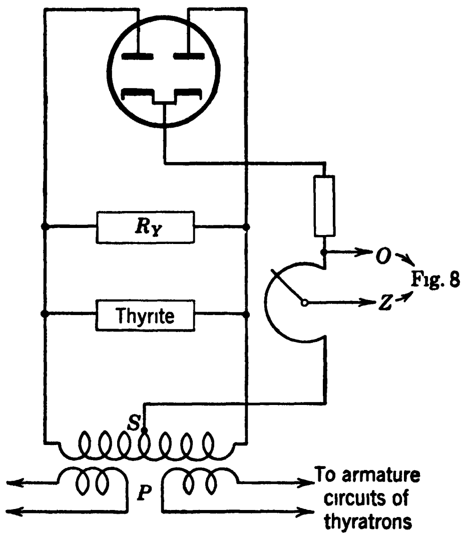


FIG. 10. Circuit for compensating for the *IR* drop in an armature.

thus making it possible for the motor to carry temporary overloads without shutdown.

The circuit of Fig. 11 will give satisfactory current limitation for a motor in operation but may fail to give the necessary protection to all circuits and parts at the instant power is applied. One satisfactory method of overcoming this initial instantaneous condition is to apply a preset grid voltage to tubes *E* and *EE* via the starting contacts on the push-button or magnetic starter. This measure assures that the initial current will be limited to, say, 150 per cent normal for the fraction of time it takes for all circuit elements to become conditioned for action.

To complete the details of the electronic control circuit for d-c motors, thyrite resistors are employed at all points where high transient voltage may be encountered. Since the action of circuits em-

ploying electron tubes is very rapid, it is possible for oscillating or hunting conditions to develop. Such hunting may be prevented by the use of filter circuits, which usually consist of a series circuit of a resistor and condenser (*RC* circuit). Such circuits are frequently placed between grid and cathode where they effect a time delay in tube action. The potentiometer type of resistor between *R*9 and *R*10 (Fig. 11) offers a fine adjustment so that the normal voltage applied to the

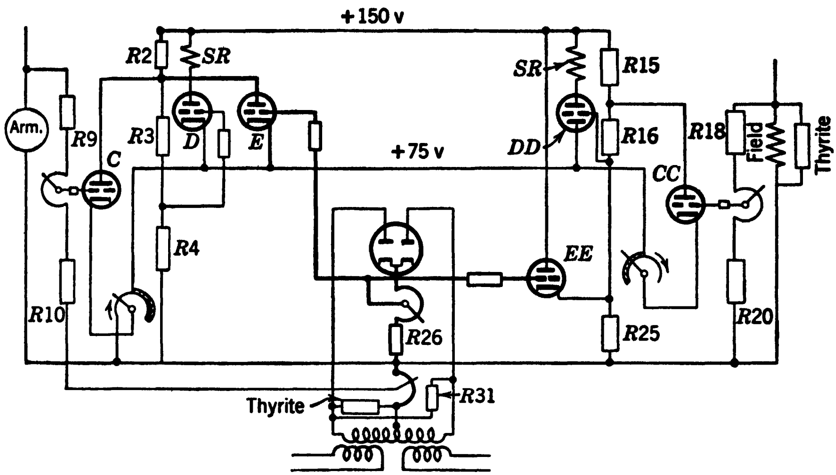


Fig. 11. Current limiter added to armature and field control circuits.

armature equals the rated value. Similarly, the variable resistor between *R*18 and *R*20 serves to give the proper field-current adjustment for normal or basic speed.

The reversing of d-c motors under electronic operation may be brought about in two ways. In one case the rectified power to the armature is removed and the armature is connected to a resistor to effect dynamic braking. After the armature comes to rest a magnetic starter may apply reversed potential to the armature to give the desired reversed direction of rotation. In case a more rapid deceleration is desired, circuits may be employed to give regenerative braking (i.e., to use the kinetic energy stored in the rotating armature to pump electric power back into the a-c system). Regeneration is not easy to accomplish under the electronic system because of the unilateral conduction of the rectifier tube. Suppose, for example, an electronic motor drive is operating a motor at three times basic speed and the field is suddenly increased to normal value. The armature counter emf

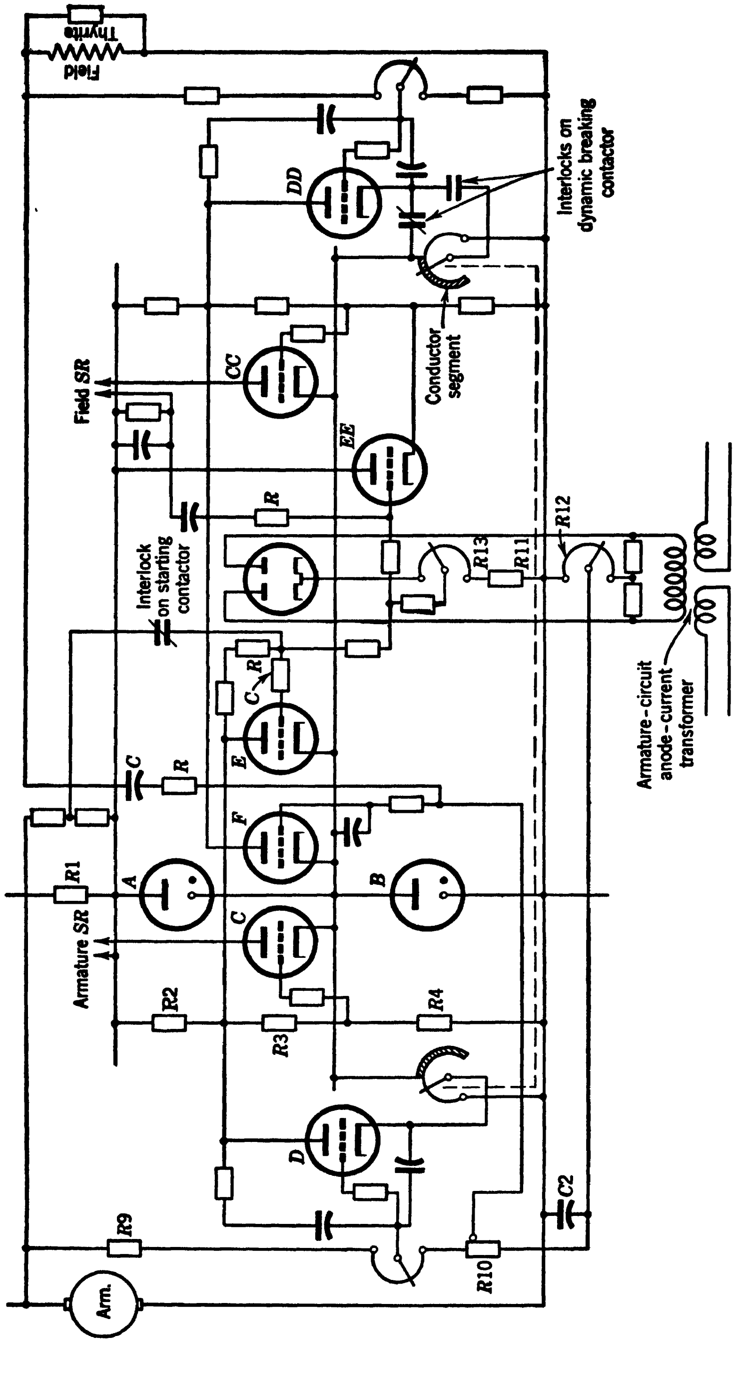


FIG. 12. Complete armature and field voltage controls for the electronic drive for a d-c motor. (Courtesy General Electric Company.)

will rise to nearly three times normal value and will be much higher than the rectified voltage of the driving thyratrons. Since the direction of electron flow cannot reverse, there is no regenerative action and the motor coasts as if the armature circuit were open. Hence in order to obtain regeneration or inverter action, the polarity of the applied

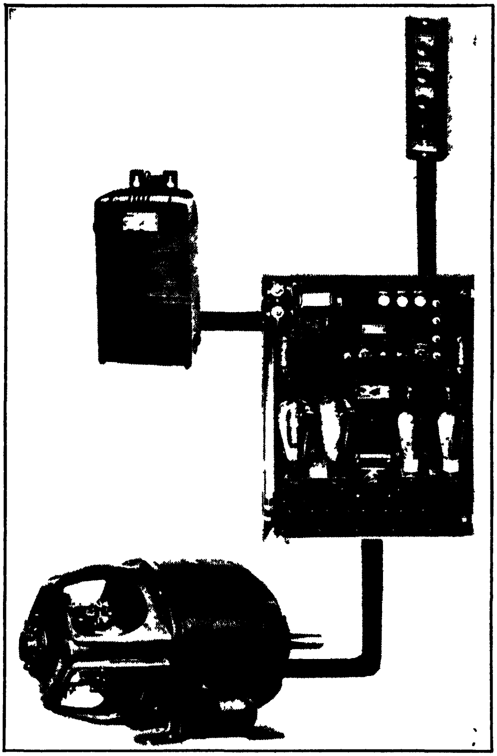


FIG. 13. Schematic assembly of the four equipment units in an electronic drive for a d-c motor. (Courtesy General Electric Company.)

voltage must be reversed by magnetic contactors. Along with this reversal the phase shifting of the grids on the thyratrons must be changed so that the regenerating action of the armature may be utilized to pump energy back into the a-c system during that short period while the motor armature is being brought to rest.

The preceding circuit illustration used a single-phase, a-c supply which is generally used for small motors. For motors of 3 horsepower and larger, a three-phase source is employed. Three rectifier tubes are required for these larger units and some changes are entailed in

the grid phasing of the thyratrons. The three-phase unit gives a smoother source of rectified supply to the armature with slightly higher over-all efficiency.

The d-c motors for electronic operation on a-c systems are built on larger frames than standard constant-speed motors because (1) the pulsating d-c current supplied by rectifier tubes causes greater heating than steady direct current, and (2) for low speeds the motor ventilation is reduced. The over-all full-load efficiency of the single-phase electronic drive is about 50 per cent. This low over-all efficiency results from the product of a conversion efficiency (single-phase) of approximately 70 per cent and a motor efficiency (small capacity and losses due to ripple) of approximately 70 per cent.

It may be of interest to note that the over-all efficiency of a small Ward Leonard motor-generator drive which gives a similar range of speed control is only 40 to 50 per cent. On units using multiphase operation the over-all efficiency will approach 60 per cent. A complete circuit for a single-phase electronic unit for d-c motor drive is illustrated in Fig. 12 and the four-unit equipment assembly for such a unit is illustrated in Fig. 13. The full-load power factor of the single-phase unit is approximately 80 per cent.

The early applications of electronic operation of d-c motors used small motors and single-phase alternating current. Applications of this system to large motors become competitive with the Ward Leonard system but the designer does not find any difficulty with the larger capacities. Three-phase power used for six- or twelve-phase rectification gives a smooth output. Ignitrons can be employed for an almost limitless current and they offer the same flexibility of control that may be secured in thyratrons. In 1947, electronic controls for d-c motors were being furnished in ratings up to 200 hp.

REFERENCE

MOYER, E. E., "Electronic Control of D-C Motors," *Electronics,* May, June, July, September, and October, 1943.

Chapter XVII

PHOTOELECTRIC CONTROL DEVICES

Many control devices utilize light as the variable medium for actuating their mechanisms. Such devices employ phototubes as detectors and some combinations of circuit elements and vacuum or gaseous electron tubes to perform the desired operations. These basic units have been combined in hundreds of different circuits to perform useful functions. Since many of the complete control circuits are somewhat complicated, it will be helpful to study some simple basic circuits before proceeding to those that are more difficult.

Basic Photoelectric Control Circuits. A simple circuit for the measurement of the amount of light falling upon the cathode of a phototube is shown in Fig. 1. This is a triode circuit in which the grid voltage is controlled by the voltage drop across the resistance R of the phototube-resistance combination. With zero light falling in the phototube, grid bias E_{cc} may be adjusted to cutoff giving zero current in the meter in the plate circuit. With light admitted to the phototube, electrons are emitted by its cathode (proportional to intensity of light) which flow through the resistor R in the direction shown and produce a rise in potential in the direction of electron flow. Thus the grid becomes less negative and an electron current flows through the plate circuit causing the meter to deflect. When operating on the linear portion of the transfer characteristic of the triode, the meter deflection will be proportional to the light intensity. The meter must be calibrated if accurate measurement of light is desired.

Simple photoelectric relay circuits for operation with d-c or a-c supply are shown in Fig. 2, parts *a* and *b*. In these circuits the indicating meter of Fig. 1 has been replaced by a magnetic relay and contactor which serve to operate a secondary circuit or equipment. Under darkness the phototube has infinite resistance and the grid of the tube is biased to cutoff. When light falls on the phototube it passes electrons through the 3-megohm resistor and varies the potential of the grid (less

negative) until the plate current is sufficient to operate the relay *CR*. When the light decreases sufficiently relay *CR* becomes de-energized. Thus this circuit gives "on" and "off" operation of relay *CR* with light variation. The sensitivity of the operation of the circuit will be determined by the grid-bias battery voltage, the resistance of the grid

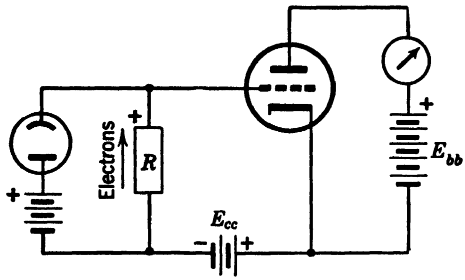


FIG. 1. Simple light-measuring circuit using a phototube.

resistor, the plate voltage, and the relay setting. The a-c circuit shown in part *b* of Fig. 2 functions in the same way as part *a*. The capacitor around relay *CR* serves to hold the relay during the half-cycles when the triode plate potential is negative.

A "lock-in" type of phototube relay circuit is shown in Fig. 3. The

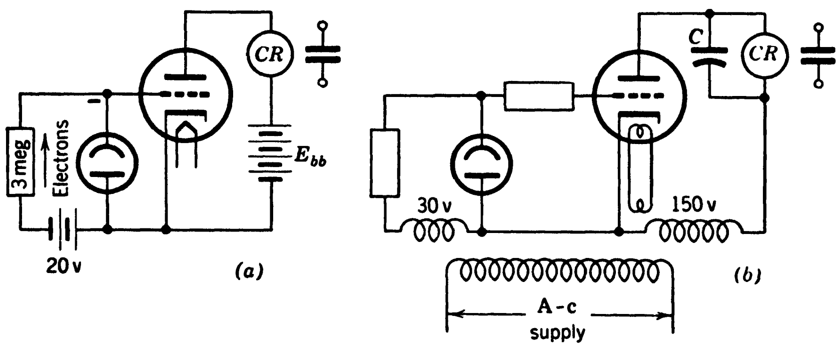


FIG. 2. Simple "on and off" relay circuits operated by light.

"lock-in" feature is secured by the substitution of a thyratron for the vacuum triode in Fig. 1. It will be recalled that, after conduction is started in a thyratron anode circuit with d-c supply, it cannot be stopped by grid control but only by an opening of the anode-cathode circuit or by a lowering of the anode-cathode voltage below the ionization potential. This "lock-in" circuit may be opened at switch *S*.

Obviously this circuit is adapted for danger signals or safety control functions.

A photoelectric relay circuit in which the relay is operated by darkness instead of by light is shown in Fig. 4. If light is falling upon the phototube and point *a* is positive the phototube will pass electrons (rectify) through the 5-megohm resistor in the direction shown. This current will produce an *IR* drop across the resistor and charge the 0.003-microfarad condenser to polarity as shown in the figure. This causes the grid of the thyratron to be negative and the tube nonconducting so that when the a-c polarity reverses with point *b* positive the relay *CR* is not actuated. During this reversal the condenser discharges slowly through the 5-megohm condenser but not fast enough to permit the thyratron to fire. When darkness falls on the phototube it ceases to rectify and the 0.003-microfarad condenser discharges after 2 or 3 cycles bringing the grid of the thyratron to the same potential as its cathode. Now the thyratron fires on each half-wave that its anode is positive and relay *CR* is energized. The by-pass condenser around *CR* stores energy for holding *CR* on negative half-cycles and resistor *R* limits the charge current into the condenser *C* to protect the cathode of the thyratron.

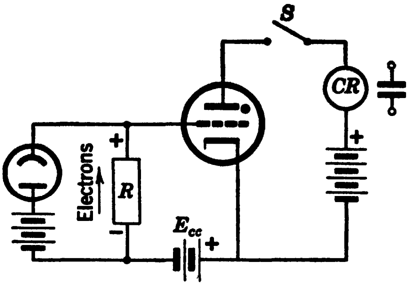


FIG. 3. Basic "lock-in" circuit using a phototube and a thyratron.

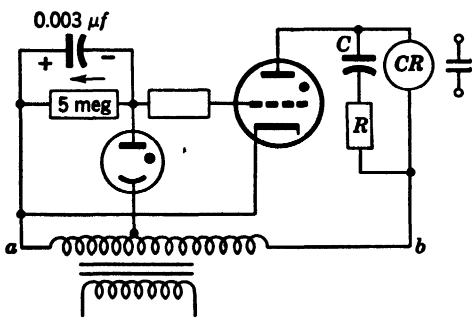


FIG. 4. Circuit holds relay closed in dark and open when light falls on phototube.

Some phototube relay devices must operate in locations that have

normal illumination subject to slow changes due to variation of daylight or other factors. These devices must have circuits designed to function on quick changes or impulses of light. Such circuits are actuated by a discharge from a condenser caused by a pulse of light. One such circuit which causes its relay to be "locked in" by a pulse of brighter (increasing) light is shown in Fig. 5. With switch S closed and normal steady illumination (or darkness) on the phototube, the grid of the thyratron is held negative by the bias battery E_{cc}. During this condition whatever light exists produces an electron flow from E_a

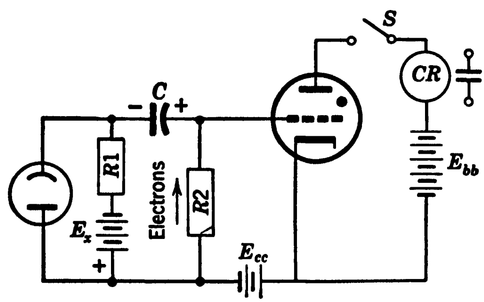


FIG. 5. Relay circuit operated by an increasing "pulse" of light.

through $R1$ and the phototube. The potential drop that exists across the phototube charges the condenser C through $R2$ to the polarity as shown. Under steady lighting conditions the charge on the condenser does not affect the grid bias on the thyratron provided by E_{cc}. Now assume that a pulse of increasing light falls on the phototube and its electron conduction rises instantly, giving a greater voltage drop across $R1$ and a fall of potential across the phototube. The fall of potential across the phototube causes the condenser C to discharge through the phototube and resistor $R2$. The discharge electrons through $R2$ give a rise (+ pulse) of voltage on the grid of the thyratron and cause it to fire. Relay CR is energized and remains "locked in" because of the d-c anode supply voltage.

A second photoelectric relay circuit which operates on a light pulse (decreasing) is given in Fig. 6. Here with switch S closed and a steady illumination on the phototube, the thyratron is held nonconducting by the grid bias battery E_{cc}. The condenser C is charged through $R2$ to the potential drop across the resistor $R1$. Now, a pulse of increasing light would raise the drop of potential across $R1$ and make the grid of the thyratron still more negative. Conversely, a pulse of de-

creasing light would lower the voltage drop across $R1$ and cause condenser C to discharge electrons through $R1$ and $R2$ in the direction shown. This pulse of electrons through $R2$ would raise the potential of the grid of the thyratron and cause it to fire and "lock in" the relay CR.

In the application of the preceding circuits, the d-c potentials indicated by a battery are usually supplied by a power pack and by voltage-regulator tubes where a constant d-c potential is desired.

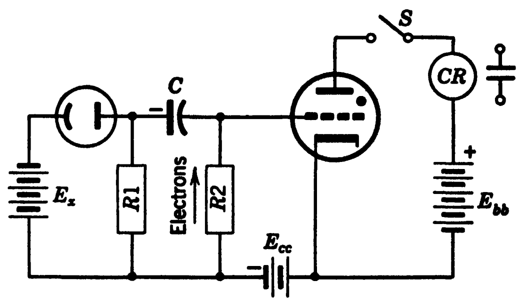


Fig. 6. Relay circuit operated by a decreasing pulse of light.

Automatic Lighting Control. Automatic lighting control is a system using photoelectric relays to turn on artificial lighting whenever daylight wanes and to turn off lighting units when they are not needed. This control may be desirable in library reading rooms, schools, offices, and factories because human beings intent on reading or working fail to notice decreasing illumination and thereby permit eyestrain and inefficiency. This system also has applications indoors for turning on emergency light systems in operating rooms in hospitals and outdoors for turning on yard lights and street lights.

An automatic lighting control system consists of a photoelectric relay unit and a magnetic contactor as illustrated in Fig. 7. The relay unit is placed in the room or location where the lighting is to be controlled, and the magnetic contactor is placed where it is convenient to open and close the lighting circuits.

The simple circuit of Fig. 2*b* will close a circuit when the illumination falls too low but the resulting rise of illumination from the artificial lighting would cause this circuit to cut out immediately and it would continue to function "on" and "off." To prevent this hunting action the circuit of Fig. 2*b* must be modified to some other form such as that shown in Fig. 8. The latter circuit will cut in the artificial

lighting at a prefixed illumination intensity and later cut out the artificial lighting when daylight rises to another prefixed level. These prefixed levels are set on potentiometers $P1$ and $P2$. In Fig. 8, $CR2$ represents the remote magnetic contactor and $CR1$ is the relay of the photoelectric relay unit. To understand the circuit operation assume that the phototube is covered with no light reaching its cathode.

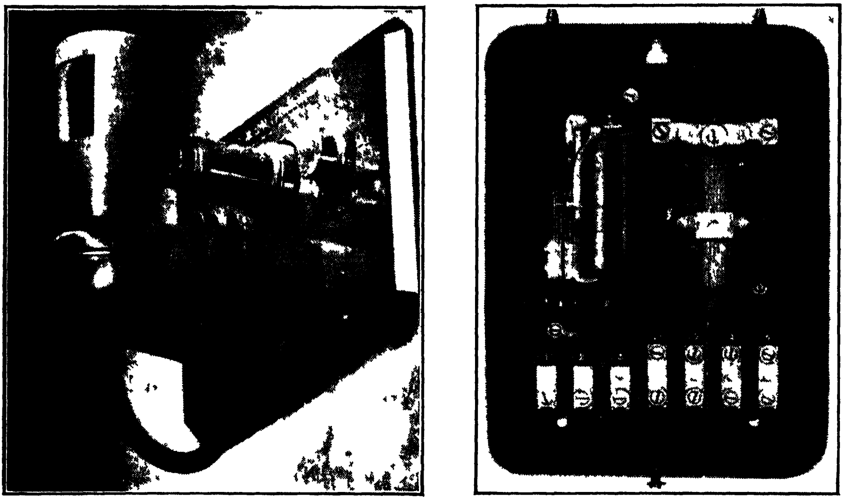


FIG. 7. Photoelectric lighting control unit with the contactor panel on right (Courtesy General Electric Company.)

Under this condition the grid of the triode will be free (no electron path through either capacitor $C1$ or the phototube), it will be negatively charged to near cut-off, and relay $CR1$ will be de-energized. The normally closed contact of $CR1$ short circuits $R3$, magnetic relay $CR2$ is energized, and artificial lights are on. In addition, contact $CR1$ between potentiometer $P1$ and the triode cathode is open. Now assume zero daylight and remove the cover from the phototube. With the proper setting of potentiometer $P2$ the triode plate current will not operate relay $CR1$. Before considering the next step, note that the phototube may be considered as an electron-leak unit for both the triode grid circuit and for the capacitor $C1$. The grid-leak circuit consists of the grid, resistor $R1$, the phototube, the lower section of $P2$ with a small a-c voltage, and resistor $R2$ to cathode. The a-c voltage across $P1$ and $P2$ is also impressed across $C1$ and the phototube in series. When the left side of condenser $C1$ is positive, the right side is

negative and electrons are withdrawn from the grid of the triode, making it less negative and permitting more electrons from its cathode to land on it. On the next half-cycle with the anode of the phototube positive, the phototube will conduct electrons from the grid toward the cathode and also will discharge the capacitor *C*1. Now assume an addition of daylight to the room and the phototube. The phototube conducts more electrons from the triode grid and discharges the condenser *C*1 more completely. As the daylight rises in value the point is reached where the triode grid voltage consisting of negative d-c

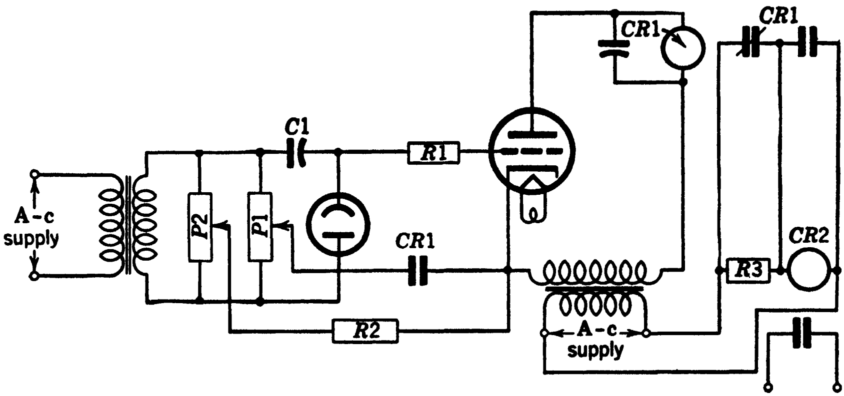


FIG. 8. Circuit for controlling room lighting.

(leak) bias plus the a-c component induced from *P*2 causes the triode to conduct sufficiently to operate relay *CR*1. The energization of *CR*1 de-energizes *CR*2 (artificial lights off) and closes contact *CR*1 between *P*1 and cathode. The closing of contact *CR*1 shifts the grid control of the triode from potentiometer *P*2 to *P*1 so as to prevent relay *CR*1 from dropping back when the artificial lights go out. When daylight wanes later to the preset point of *P*1 relay *CR*1 will drop out and the artificial lighting will be re-established.

Photoelectric Counting and Operational Devices. The most common application of photoelectric relays is for counting passing objects and for initiating the operation of various kinds of equipment. These functions are actuated by the interception of a beam of transmitted light. The beam of light is produced by an enclosed source of artificial light and is directed upon a phototube in a photoelectric relay circuit as illustrated in Fig. 9. Any opaque object passing through the light beam serves to actuate the photoelectric relay and its associated circuits. An inexpensive form of photoelectric relay

known as a Photo-Troller which is suitable for light interception operation is shown in Fig. 10 and a light-source unit is depicted in Fig. 11. This Photo-Troller is a relatively insensitive device which requires an illumination of 40 foot-candles or more for operation. It may be

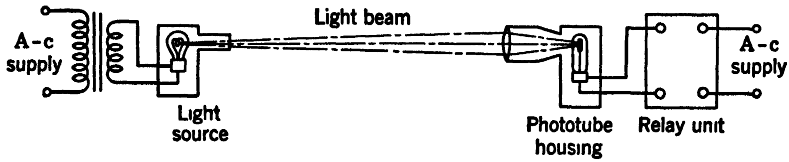


FIG. 9. Electronic control via a light beam and phototube.

used in a location having normal light provided that the illumination on the cathode of the phototube from this source is less than 40 foot-candles. The circuit of the Photo-Troller is designed to energize its relay coil when the phototube is dark and to de-energize the relay when the phototube is lighted, the predominating condition encoun-

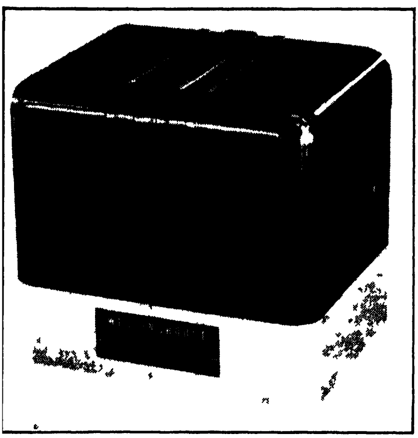


FIG 10. Photo-Troller, a photoelectric relay. (Courtesy Westinghouse Electric Corporation)

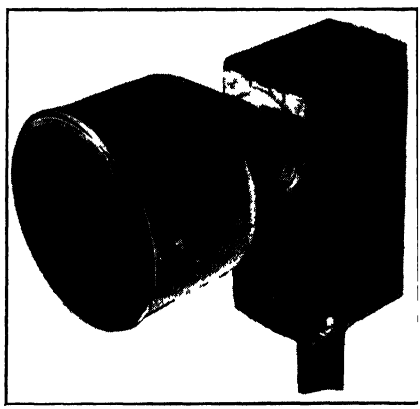


FIG. 11. Light source for photoelectric control (Courtesy Westinghouse Electric Corporation)

tered with simple photoelectric applications. The circuit shown in Fig. 12 employs a gaseous phototube and a shield-grid thyratron. The firing of the thyratron is controlled by two voltages applied in series in its grid circuit. One of these voltages is an a-c voltage drop across resistor $R3$. The second of these voltages is a d-c negative bias developed across resistor $R2$ and capacitor $C1$ by the rectifying action of

the circuit from the pointer on $P1$, the phototube, resistor $R2$, and resistor $R3$. It should be noted that $R3$ and $C2$ in series across a 30-volt alternating current constitute a phase-shift circuit which shifts the phase of the a-c drop across $R3$ ahead of the a-c voltage across the anode-cathode of the thyratron. This shift insures prompt and early firing of the thyratron. The magnitude of the negative d-c bias across

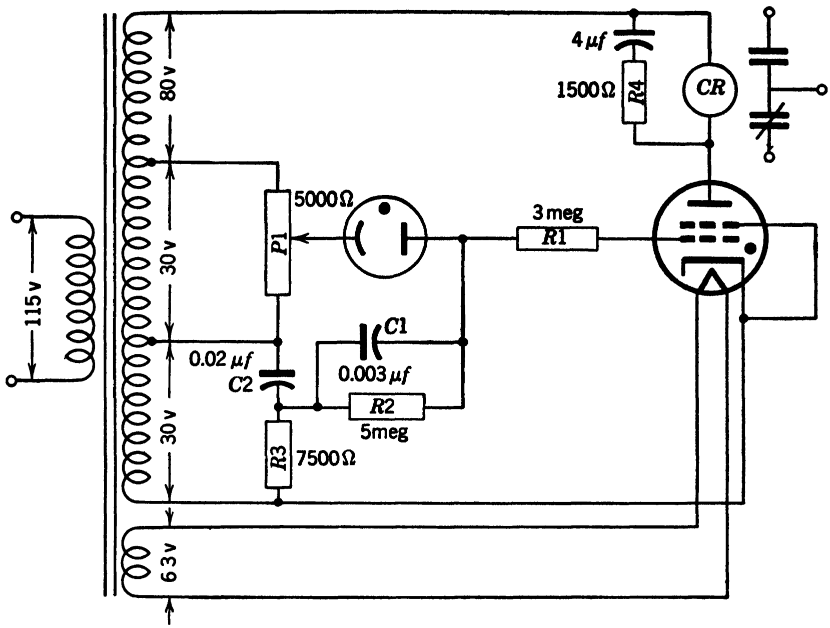


FIG. 12. Circuit for a photoelectric relay.

$R2$ is determined by the setting of $P1$ and the illumination on the phototube. For illumination above 40 foot-candles (approximately) the d-c negative bias will prevent the thyratron from firing. When the light beam is intercepted the illumination on the phototube falls and the negative bias decreases and permits the a-c voltage across $R3$ to fire the thyratron on half-cycles when its anode is positive. Restoration of light stops the thyratron conduction and de-energizes the relay CR.

If the contactor on relay CR of Fig. 12 is connected to a magnetic counting device, the system may be used for counting parts on a moving conveyer system, packages on a belt conveyer, automobile traffic on a highway, or customers entering a business establishment. Similar association of the relay contactor with a door-opening mech-

anism provides automatic opening of garage doors for automobiles, warehouse doors for trucks or trains, and doors for pedestrians (see Fig. 13). Another application of this system is to open the valve on a drinking fountain when a person bending over to drink intercepts a beam of light.

Sorting and Inspection of Photoelectric Relays. Sorting and inspection operations by photoelectric relays usually employ reflected or transmitted light wherein the light-detection mechanisms must be sensitive and discriminating. Sensitivity in a photoelectric relay circuit requires the use of an amplifying stage between the phototube and the trigger tube, and light discrimination may require that illumination from external sources be eliminated. One form of a sensitive photoelectric relay circuit is shown in Fig. 14. In this circuit the firing of the shield-grid thyratron on the right is controlled by two separate voltage sources applied in its grid circuit. One voltage is the a-c phase-shift voltage across capacitor $C2$ and the other is a negative d-c bias across points XY which is controlled by the light falling on the phototube. The amplifying circuit consists of a pentode having d-c potentials applied to its electrodes by batteries or a power pack. The variable grid bias of the pentode is provided by potentiometer $P1$ and the input-grid voltage is provided by the voltage drop across resistor $R1$. This grid-input voltage is controlled by the light falling on the phototube. The voltage across XY is the difference between the fixed battery potential E_x and the IR drop across $R2$ due to the current in the pentode plate circuit. With the double-pole double-throw switch S thrown to the right, a rise in illumination on the phototube will make the grid of the pentode less negative, increasing its

Fig. 13. Photoelectric automatic door opener. (Courtesy General Electric Company.)

plate current and thus decreasing the negative potential on the thyratron. At a predetermined point the thyratron will fire and operate relay *CR*. With the reversing switch *S* thrown to the left and with normal illumination on the phototube the potentiometer *P*1 is adjusted until the drop across *R*2 exceeds the battery potential E_x, making the thyratron grid negative and nonconducting. With these adjustments the thyratron will be fired by a decrease of light instead of an increase.

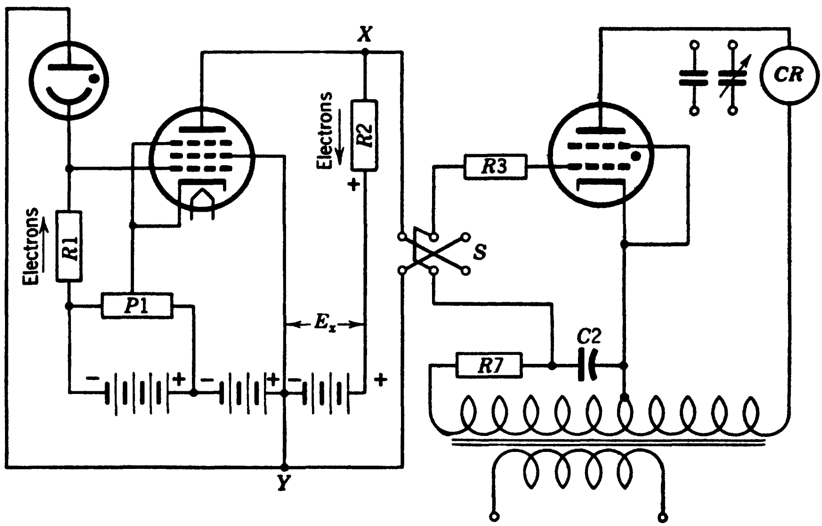


FIG. 14. Sensitive photoelectric relay circuit.

Sorting operations require the use of some form of scanning device wherein light thrown upon objects on a moving belt is reflected into a photoelectric relay head. One form of scanning head is shown in Fig. 15. Seeds such as beans and corn have been sorted successfully. Those seeds or objects which are lighter or darker in color than the standard cause the photoelectric relay to operate and this operation causes an electromagnetic device to reject the defective unit. Packaged articles moving on a belt can be checked for absence of labels. Eggs have been candled with some degree of success by measuring absorption of transmitted light by photoelectric devices.

Steel sheets intended for tin plating are inspected for the detection of holes by passing the plate through the photoelectric scanner illustrated in Fig. 16. A bank of artificial lights plus a blower for cooling is located in the upper unit and a row of phototubes in the lower unit detects any light coming through holes in the plate. Holes as small as

0.015 inch are detected at speeds up to 100 feet per minute. The operation of the photoelectric relay due to a pinhole defect may cause the sheet to be diverted into a separate channel or may control a mechanism that marks the plate by means of a knurled roller.

The efficiency of combustion in a furnace may be determined by a photoelectric scanning of light transmitted through the flue gases. Here the photoelectric relay rings a bell alarm when the smoke density becomes too great or the relay action can be made to control the rate of air feed to the furnace and thus regulate the combustion. In a similar manner the turbidity of water or the density of a solution may be inspected by the transmission of light to a photoelectric relay.

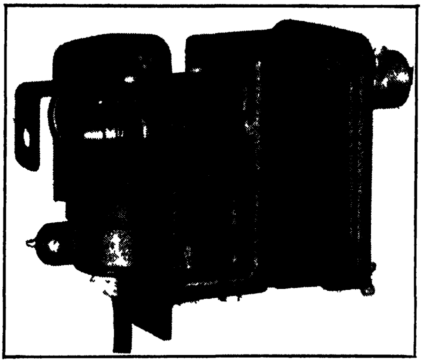


FIG. 15. Photoelectric scanning head. (Courtesy General Electric Company.)

Photoelectric Protection Devices. Many photoelectric devices are used to protect human beings, machines, and property. One such device serves to cut off the fuel supply to a burner when the flame in the furnace is extinguished. A relay circuit similar to the basic circuit of Fig. 4 may be employed although a more sensitive and hence more complicated one may be desirable.* Dangerous working regions near the jaws of a punch press and moving parts of a machine may be guarded by a beam of light which stops the machine when the human hand is in danger of injury. In like manner, beams of light may serve to prevent the starting of a machine when obstructions are present that would cause damage to the machine itself. Banks, stores, and factories may be guarded at night by beams of invisible light (infrared or ultraviolet) which when intercepted by intruders set off alarms through photoelectric relays. Many war plants were guarded against sabotage in this manner. Photoelectric relays are often used as limit switches to stop assembly-line belts and conveyers when material becomes congested at some plant near the end of the line. The breakage of a paper sheet, a wire, or a thread on a roll feeding a production machine may serve to shut

* Nonphotoelectric systems using thermostatic units or the conductivity of a flame itself are often employed for flame protection.

down the machine through photoelectric action. Applications of photoelectric protection devices are too numerous to list.

Photoelectric Register Controls. Photoelectric register controls serve as guides for paper or metal passing rapidly in strip form over rolls. Registor controls are divided into two classifications, cutoff and side register. The side register and correction register are illustrated in Fig. 17. In this register the material and roller are of contrasting colors and the control is arranged so that, with the roller

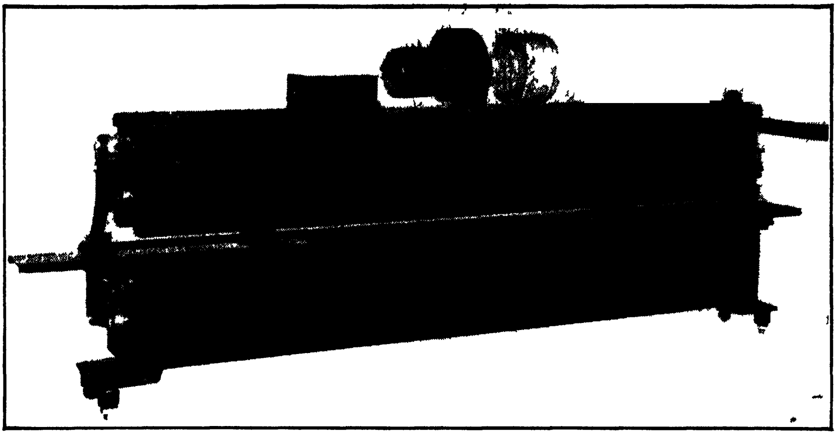


FIG. 16. Photoelectric scanner for the detection of holes in tin plate. (Courtesy General Electric Company.)

under the scanning head, the feed and idler roll are moved to the right. If the material is under the scanning head, the feed and idler rolls move to the left. Thus a balance is reached that will keep the edge of the material under the middle of the light spot. With this equipment it is possible to hold the edge of the material to $+\frac{1}{64}$ inch while it is traveling at 1500 feet per minute.

The cutoff register is operated by a register mark or spot printed on the material being controlled. The passing of this register mark creates an impulse in the phototube circuit as it passes the scanning point. This impulse fires the thyratron and operates a mechanical knife for cutting the material, or it may serve as a corrector of the material speed so as to cut the material in accordance with printed forms on paper used for wrapping.

A form of photoelectric inspection equipment known as a weft straightener control is used in the textile industry. In this application, phototubes scan the cloth as it passes through a machine and de-

tect any skew that is present in the weft threads. When skew is present a difference in the frequency impulse received by two phototubes is noted and this difference in frequency serves as a medium for correcting the skew in the weft threads.

Register control is also used on printing presses for multicolor work to insure the proper alignment or match of the different colors.* Normally, alignment is difficult since paper may change its dimensions as much as 1 per cent because of changes in humidity, tension, temperature, and other factors. A schematic drawing of the equipment used is shown in Fig. 18. Two photoelectric scanning heads are used. One scans slits in a ring or disk mounted on the impression or printing cylinder (see lower right of figure); the other scans register marks along the margin of the paper (web) which are printed by the first color process (usually yellow). The impulses received by the two photoelectric scanners are fed to a comparator mixing panel and circuit which utilizes any time difference to vary the angle between the indrive shaft and the impression roll so that the successive color design covers properly the first and succeeding color impressions. It is claimed that this equipment will hold the later color impressions within an accuracy of 0.002 to 0.003 inch.

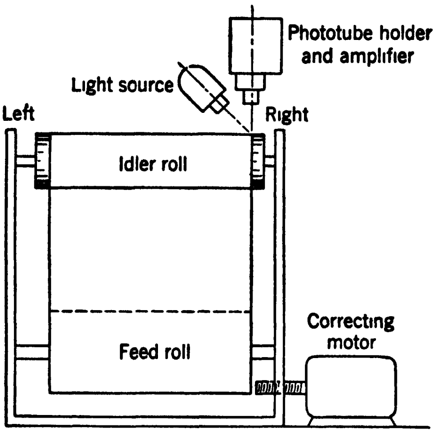


FIG. 17. Diagram showing the scanning and correction method used in a side register control on a paper-handling machine.

Photoelectric Color Selection and Analysis. Objects of different colors may be selected and separated photoelectrically by two methods. First, if the objects of a given color are of uniform shade and if the reflection or transmission of the colors represented differ sufficiently in intensity, a photoelectric device can be designed to distinguish between them satisfactorily. This would be a rough form of differentiation. A second and more exact type of qualitative selection can be secured by use of several photoelectric relays, each equipped with a different color filter.

* See article by W. D. Cockrell on photoelectric register control for multicolor presses in *Printing Equipment Engineer,* August 1941.

In material testing, production, and research, it is often desirable to match colors or to make an analysis of the color of a given material. The analysis of the color of an object consists in determining the percentage of light reflected or transmitted (in materials such as filters or solutions) by the object at each wavelength in the visible spectrum. Thus, the sample to be measured is illuminated with spectrum light of substantially a single color, and the light reflected or transmitted by the sample relative to a standard color is plotted against the wave-

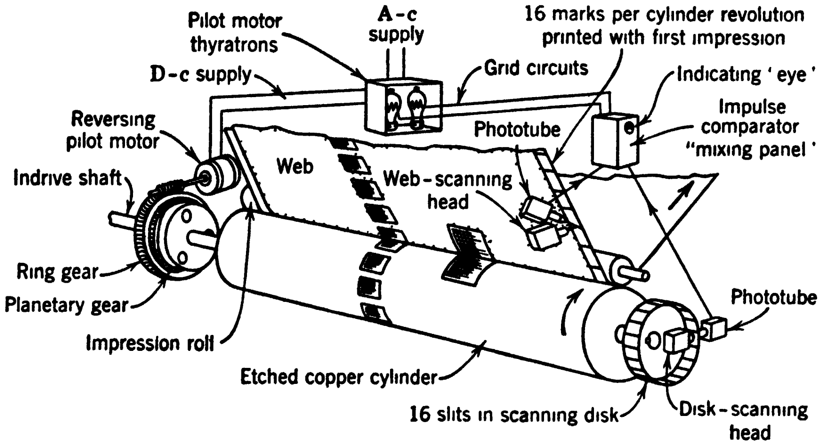


FIG 18 Automatic register control for rotary webb presses showing the general location of parts (Reprinted from *Printing Equipment Engineer,* August 1941)

length of the incident light. The standard color—white—which forms a basis for color comparison, is a freshly smoked magnesium oxide surface, chosen because of its uniform reflectance of practically all the incident light in the visible spectrum. This white color has been accepted internationally as the standard for color measurements.

Color analysis can be performed in a laboratory by the standard bar photometer method. This method requires hundreds of readings and requires hours for taking readings, making calculations, and plotting the desired color-spectrum chart. A complete analysis of the color of an object including the rendering of a spectrophotometric curve can be performed automatically in a few minutes by the use of a phototube and appropriate auxiliary equipment. A commercial design of an instrument for performing this service is shown in Fig. 19. In this device white light is passed through a prism to separate it into the color spectrum. Then the different colors in this spectrum are passed to the

object, the resulting reflection or transmission is compared with the standard, and the ratio is printed on a chart. Commercially the device is very useful for matching colors of textiles, paints, and plastics.

Photoelectric Pyrometer. The photoelectric pyrometer is a radiant-energy-responsive device for indicating or recording the temperature of incandescent bodies. The radiant energy from a hot body is directed to a phototube and causes it to pass a current which bears a definite relation to the temperature of the hot body. This current is

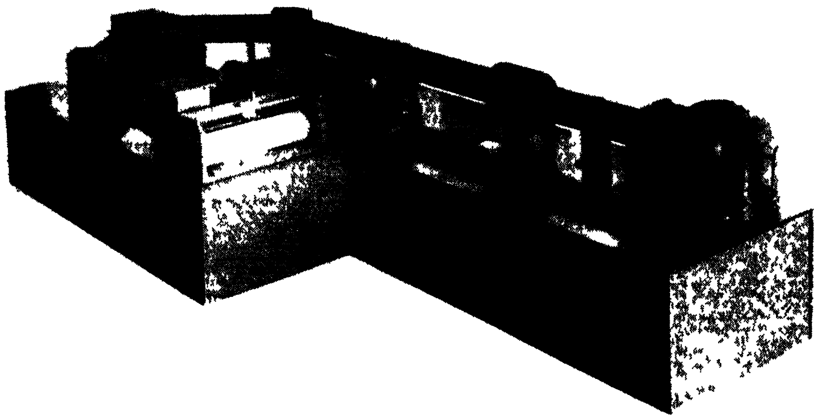


FIG. 19. Recording spectrophotometer. (Courtesy General Electric Company.)

amplified by an extremely stable vacuum-tube amplifier and the amplified current is passed through an indicating or a recording instrument. The principal feature of this pyrometer is that it provides, without appreciable time lag, a continuous indication or record of the temperature of incandescent bodies.

A heated metal as it approaches its melting point passes through progressive stages in which it has an initial color of red which changes to orange, yellow, and finally white. With an increase in temperature the amount of radiant energy increases exponentially. Phototubes are sensitive to radiant energy throughout the visible spectrum though the response varies as discussed in Chapter VII. The relation between the phototube response and the temperature is a very steep one in the spectrum range for incandescent bodies. For example, the phototube response to a temperature of 2150 degrees F is approximately ten times that for a temperature of 1700 degrees F, or, for a smaller temperature range, a change from 1800 degrees F to 1850

degrees F causes an increase in phototube response of 32.8 per cent. From this relationship it is obvious that a milliammeter could be calibrated to indicate the temperature of an incandescent body provided that a stable circuit and amplifier system is used. A vacuum-type

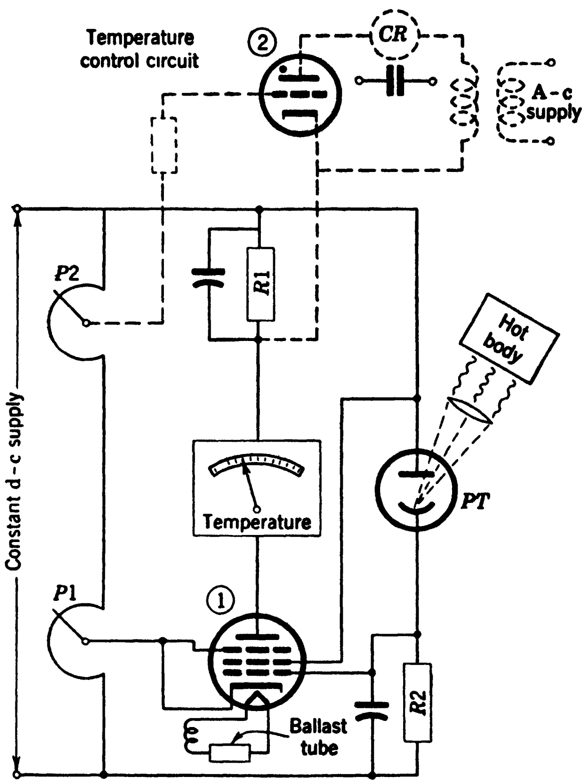


FIG. 20. Schematic diagram of a phototube pyrometer circuit and control equipment.

phototube with a caesium-oxygen-silver cathode is employed to utilize through a suitable optical system the radiant energy from a specific area of the incandescent body.

A photoelectric pyrometer may employ the schematic circuit shown in the full-line diagram of Fig. 20. A constant voltage applied on the left of the diagram is provided by the conventional power pack plus a voltage-regulator tube circuit. A phototube *PT* and a resistor *R*2 act as a potential divider to control the bias on the control grid of the pentode tube 1. The potentiometer *P*1 is adjusted to hold the cathode

of the pentode at a higher potential than its grid. As the radiant energy from the hot body rises, the phototube conducts an increasing current which raises the drop across $R2$ and the potential of the control grid of the pentode. This rise in grid potential controls the cathode-plate current of the pentode which, in turn, operates the temperature-indicating milliammeter. The temperature-indicating meter may

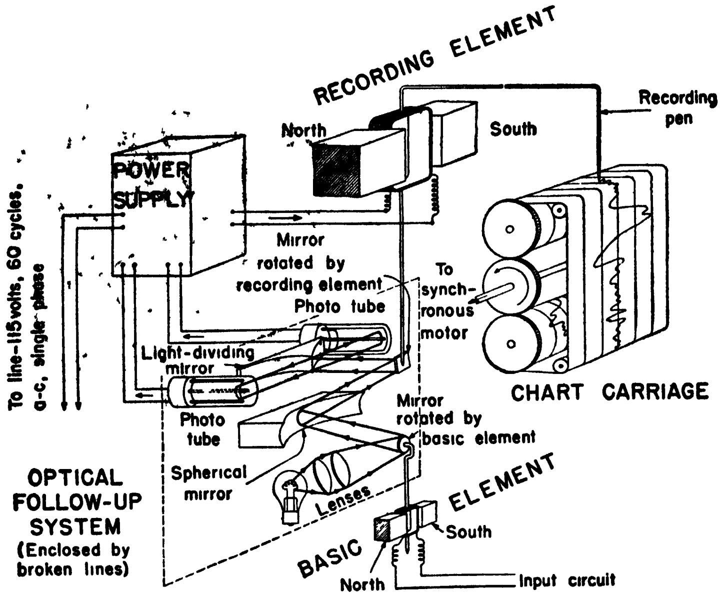


FIG. 21. Functional construction and circuit of a photoelectric recorder. (Courtesy General Electric Company.)

be placed at a distance of 50 feet from the control panel. If a record of the temperature is desired, a graphic recording meter may be substituted for the indicating meter.

A production process involving incandescent metals may require an audible signal when a certain temperature is reached or it may require some control of action to take place when one or more temperatures are reached. Such control may be accomplished by the addition of the temperature-control circuit (dotted) of Fig. 20. This additional circuit employs a thyratron 2 and a control relay CR in an a-c supply circuit. The potential of the cathode of the thyratron is determined by the drop of potential across resistor $R1$ while the grid potential is governed by the setting of potentiometer $P2$. Thus $P2$ is set to fire the thyratron at the desired temperature. If control operations are de-

sired at more than one temperature, more control circuits like the dotted lines of Fig. 20 are added and potentiometers corresponding to $P2$ are set for the desired temperatures of action.

Photoelectric pyrometers are used in steel mills to indicate the temperature of steel billets before they enter the process for making seamless tubes, lap welding steel tubes, or rolling structural shapes. The proper temperature of the steel billets coming from different furnaces assures a more uniform product with less spoilage. Similarly, a photoelectric pyrometer control has been used on a spinning machine for making cast-iron pipe by the centrifugal sand-spun process. The pyrometer indicates the temperature inside of the pipe and a supplementary control stops the spinning motor when the proper temperature is reached. In another application the temperature of the cement clinker in the kiln of a cement mill is indicated and auxiliary controls regulate the temperature by governing the speed of rotation of the kiln.

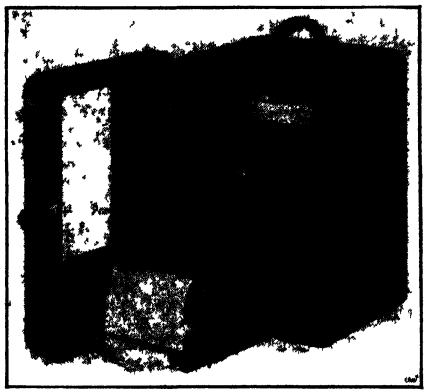


FIG. 22. Photoelectric recorder. (Courtesy General Electric Company.)

Photoelectric Recorder. The photoelectric recorder is a high-speed high-sensitivity graphical recording instrument which utilizes light and phototubes for its response. The instrument may be used for recording d-c current, d-c voltage, frequency, telemetering, or illumination. The principle of operation and the functional relation of the parts of the device are illustrated in Fig. 21.

The shafts of the recording and measuring elements, each carrying a small mirror, are on the same vertical line. The light beam is provided by the small incandescent lamp, the light from which is focused by the pair of condensing lenses. This beam is reflected, in turn, by the basic mirror, the spherical mirror, and the recording-element mirror, until it is finally divided between the two phototubes.

When the mirrors of the recording and the basic elements are parallel, the light is equally divided between the two phototubes, and no current flows to the recording-element coil. If the pen is slightly to the right of its correct position, a greater part of the light falls on the nearer phototube. This causes a current to flow from the power supply through the coil in a direction that moves the pen to the left.

Errors are reduced to negligible amounts because the function of the optical-balancing system depends upon division of the light rays, and not upon the light intensity. Also, the slightest unbalance causes the power unit to give full output of current. All the tube characteristics can change widely without affecting the final result.

A commercial form of a photoelectric recorder is shown in Fig. 22.

Problems

1. In the circuit of Fig. 1 the following components are used: triode of Fig. 11, Chapter IV; vacuum phototube of Figs. 4 and 7 of Chapter VII; grid bias = -12.5 volts, plate voltage = 200 volts, and $R = 5$ megohms. When 0.1 lumen falls on the cathode of the phototube, what is the reading on the milliammeter? What is the current in the phototube? What is the current gain in the circuit?

2. If the gaseous phototube of Fig. 8, Chapter VII, is used in part *a* of Fig. 2 with 0.2 lumen on its cathode, what will be the grid bias on the triode?

Chapter XVIII

ELECTRONIC REGULATORS

Voltage Regulation. The terminal voltage of a generator under load may be maintained constant by a suitable control of its field current. The field current may be controlled (1) on a constant-supply voltage system by a variation of the resistance in series with the field, or (2) by a direct variation of the voltage supply to the field. The first method is used in the constant-voltage system in the modern automobile or truck where vibrating relays vary the resistance in series with the car generator. Another form of method (1) utilizes an electromagnetically controlled variable resistor in series with the field. The electromagnet is bridged across the voltage to be regulated and gives a pull proportional to the voltage. Changes in voltage produce a change in magnetic pull which operates spring-mounted contacts that cut "in" or "out" resistance units in such a way as to counteract the voltage change producing the movement. The second method of voltage regulation, which varies the current supplied to the field, employs electronic servomechanisms.

The basic principle of electronic voltage regulation of a generator is to balance a constant reference voltage against the voltage to be regulated and to utilize any existing difference of potential for controlling the return to normal. A simple circuit for regulating the voltage of a small-capacity d-c generator or an exciter is given in Fig. 1. A constant reference voltage RV in series with the terminal voltage across G is applied between the cathode and grid of the triode which controls the current to the generator field F. Assume a desired terminal voltage of 100 with P adjusted to give an RV of 100. Let the generator voltage rise to 101. Then, tracing potentials from c to g will give

$$+100 - 101 = -1$$

and the resulting negative potential of 1 volt applied to the grid will reduce the current through the triode to the field and bring the generated voltage back to normal. A similar simple electronic control ap-

plied to a three-phase alternator is shown in Fig. 2. Here the terminal voltage from one phase of the alternator is rectified in a vacuum double diode and balanced against a d-c reference voltage *RV*. It should be noted that the rectified voltage is an average value which may not bear a constant ratio to the effective value under conditions

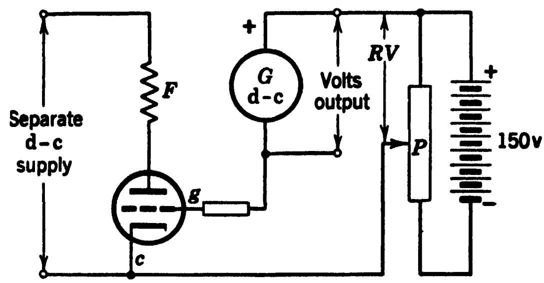


FIG. 1. Simple balanced circuit for the regulation of the voltage of a d-c generator.

where the load on the alternator may change the wave form of the terminal a-c voltage.

The magnitude of the field current required for d-c generators and exciters is usually beyond the carrying capacity of vacuum amplifier tubes so that some other form of amplification must be included in the

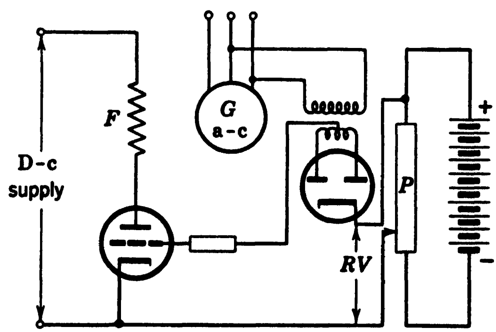


FIG. 2. Circuit for the voltage regulation of an alternator.

servo regulating system. An electronic amplifier for the field current is illustrated in Fig. 3. In this circuit the error-detecting triode supplies direct current to a saturable reactor in a phase-shifting unit which, in turn, controls the grids of a full-wave thyratron rectifier. A second method of amplifying the error-detector signal is to apply this

signal to the field of a rotary amplifier such as an amplidyne, Rototrol, or Regulex (Fig. 4). In some cases a sensitive d-c exciter may be used in this circuit. These rotary amplifiers are certain to give a time delay

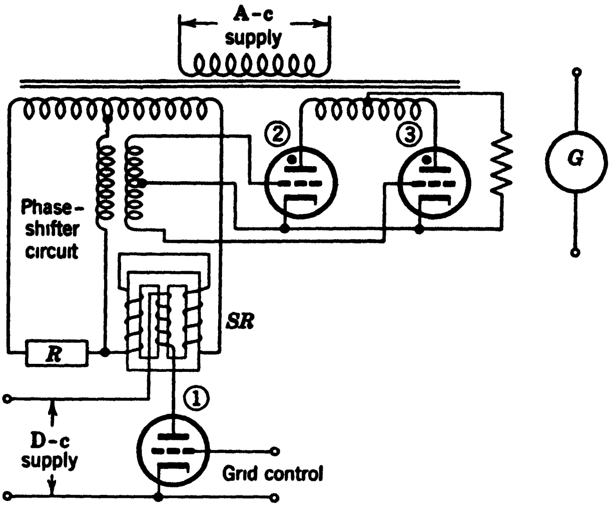


FIG. 3. Voltage regulation using thyratrons and a saturable reactor.

in the correcting action because of the inductance in various parts of the closed circuit. This time delay usually results in overshooting and a hunting action. Thus antihunt circuits are necessary and one of these is shown in Fig. 6.

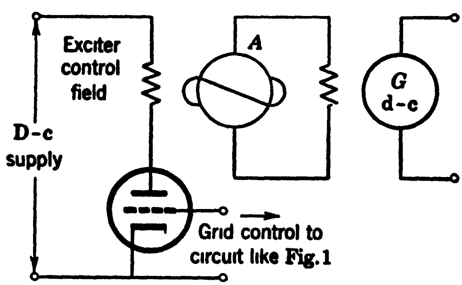


FIG. 4. Voltage regulation with an amplidyne.

A commercial form of an electronic voltage regulator is illustrated in Fig. 5 and a detailed circuit for its operation and application is given in Fig. 6. Although Fig. 6 appears complicated, it employs merely an

extension of the principles shown in Figs. 1, 2, and 3. All the circuit to the left of line *CL* is contained within the regulator. Three-phase alternating current is employed for the source of power. Transformer *TT*1 furnishes current for heating the cathodes of five tubes. Transformer *TT*2 is a power unit for a three-phase thyratron rectifier consisting of tubes *T*1, *T*2, and *T*3. Transformer *TT*3 is a three-phase grid transformer for firing the three thyratrons cited. Transformer *TT*4 supplies power to a three-phase Rectox (blocking-layer) rectifier and filter which, in turn, furnishes a d-c plate supply for the voltage comparison tetrodes *T*4 and *T*5. These tubes are connected in parallel so that, if one should fail, the other will continue to regulate. Similarly, if one of the three thyratrons should fail, the other two will continue regulation of voltage.

FIG. 5. Electronic voltage regulator. (Courtesy Westinghouse Electric Corporation.)

The full-line circuits to the right of line *CL* contain the additional equipment necessary for regulating the voltage of a d-c generator having its field supplied by a separate exciter. Following the theory of Figs. 1, 2, and 3, the d-c generator terminal voltage appears across resistors *R*16 and *R*17 and the reference voltage *RV* equals that across battery E_{cc} plus the drop across resistor *R*7, the latter being a portion of the potential divider *R*5-*R*6-*R*7. The difference voltage appears in the cathode-grid circuit of tubes *T*4 and *T*5. This circuit can be traced from cathodes, lower line, resistor *R*7, *R*16, *R*17, *R*15, battery E_{cc}, *R*14, *R*12 or *R*13, to grid. The cathode-to-grid voltage of the thyratrons consists of a d-c component and an a-c component in series. The d-c component is supplied by the voltage drops across resistors *R*4 and *R*5, while the a-c component is supplied by the grid transformer *TT*3. The a-c component, 16 volts in magnitude, is shifted 90 degrees with respect to the a-c anode potentials.

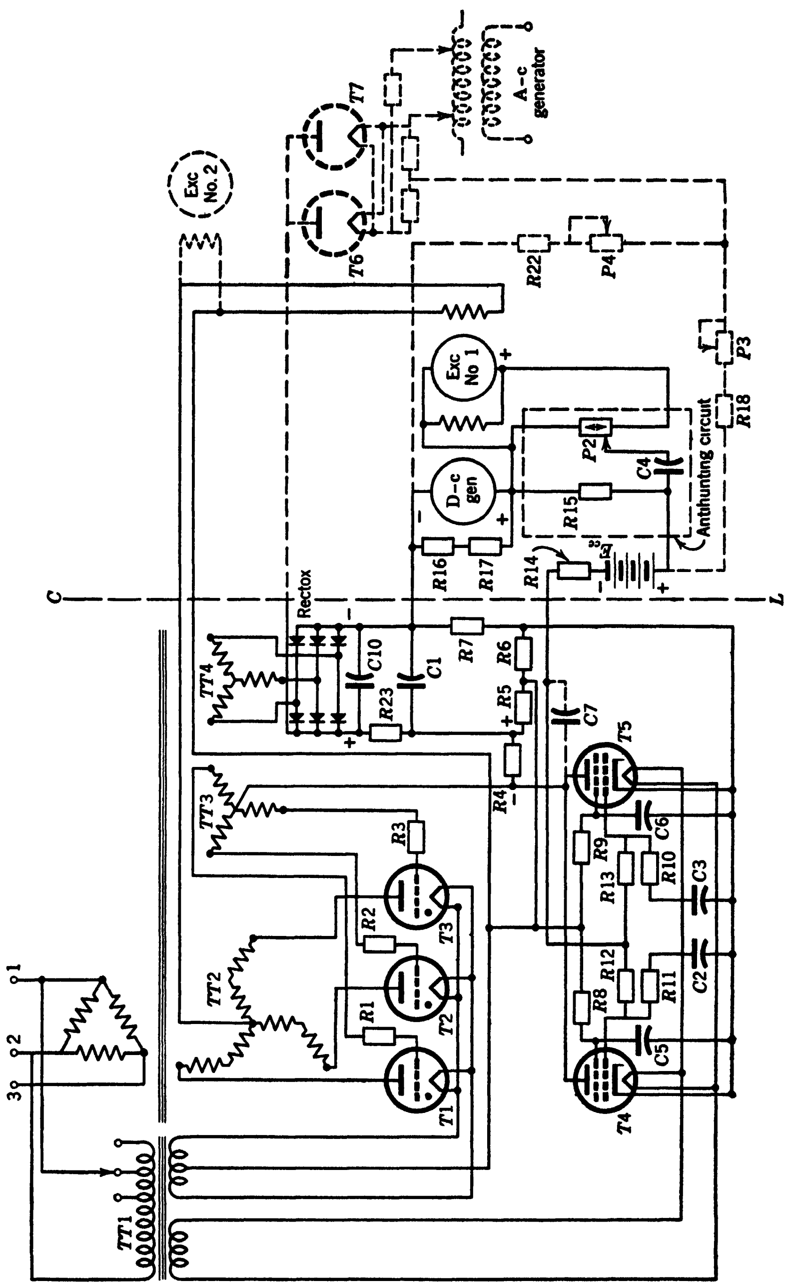


FIG. 6. Circuit for an electronic voltage regulator.

To follow the chain of regulating action, assume that the terminal voltage of the d-c generator falls below the desired value. This change will make the grids of the regulating tubes more negative and will reduce the current conducted in their cathode-plate circuits. This reduction of current flowing through resistor $R4$ will reduce its RI drop and lower the negative d-c bias on the thyratrons. Now the thyratrons will fire earlier, conduct more current, increase the field current to exciter No. 1, increase its generated voltage, increase the current through the generator field, and raise the generator voltage.

The sequence of actions given in the last sentence will result in a time delay which will cause the generator voltage rise to "overshoot," followed by a reversed action with subsequent oscillations called hunting. Such hunting can be controlled by the antihunting circuit shown within the dotted lines on Fig. 6. It should be noted that resistor $P2$ is connected across exciter No. 1 and hence carries a current proportional to its terminal voltage. Resistor $R15$ lies in the path of the cathode-grid circuit of the regulating tubes $T4$ and $T5$. From the sequence of actions suggested at the end of the preceding paragraph, it is evident that a rise of voltage will occur across the exciter and $P2$ before the generator voltage does rise. Since $P2$ and a capacitor are in parallel with $R15$, the rise of voltage across $P2$ in the direction of the electron flow (arrow) will introduce a voltage across $R15$ before the generator voltage starts to rise. This transient voltage across $R15$ is in such a direction as to make more positive the control grids of tubes $T4$ and $T5$. Thus that voltage difference which tends to raise the generator voltage is snubbed in the sequence chain before the rise actually reaches the generator terminals. This snubbing action is restrained by the RC time constant of the antihunt circuit so that satisfactory regulation occurs but hunting is prevented.

The voltage regulator of Fig. 5 can be adapted to control the voltage of an alternator by using the circuits and equipment shown by the dotted lines on Fig. 6. Here a novel method is employed for balancing the a-c voltage against the reference voltage. This method employs two vacuum diodes $T6$ and $T7$ connected across the three-phase power pack via resistors $R22$ and $P4$. The filamentary cathodes are energized through a filament transformer connected to the a-c voltage to be regulated. Thus the cathode current and voltage (also heating) are proportional to the effective value of the a-c voltage. If the diodes are operated on the linear portion of the cathode-current plate-current curve (Fig. 11, page 49), the diode should conduct a current proportional to the a-c effective voltage. This latter current is passed

through resistors $R22$ and $P4$, giving a voltage for balancing against a reference voltage as in the preceding example. In this application the thyratrons control the field current of exciter No. 2 for the alternator (not shown). Hunting is not likely to occur but if it does it can be snubbed by the addition of condenser $C7$.

A large number of electronic circuits for voltage regulation are in use. Most of them utilize the basic principles shown in Figs. 1, 2, 3, and 4.

Motor Speed Regulation. The speed regulation of a d-c motor can be effected by a slight modification of the principles and equipment

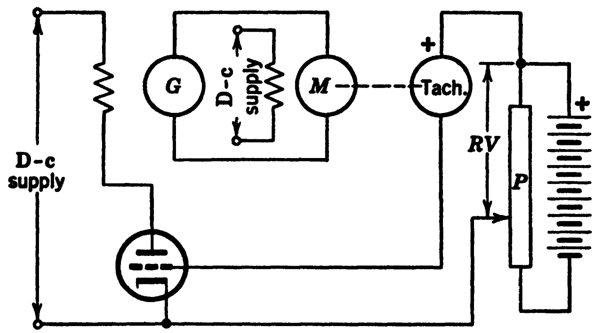


FIG. 7. Schematic circuit for motor-speed regulation.

discussed in the preceding article. To apply the principle of voltage regulation, a pilot tachometer generator having a constant permanent-magnet field should be direct-connected to the motor to be regulated. Such a pilot generator delivers a voltage directly proportional to speed. Next, the motor may be driven by the Ward Leonard system of speed control where its field is energized by a constant d-c voltage source, while the armature is connected to a separately excited d-c generator. A simple application of speed regulation is shown in Fig. 7. Here, as in preceding servomechanism regulating circuits, the terminal voltage across the tachometer (proportional to speed) is balanced against a constant reference voltage RV and the difference voltage applied to the cathode-grid circuit of a triode. The control tube regulates the voltage produced by the generator which, in turn, controls the speed of the motor (speed directly proportional to armature voltage when field flux is constant). Various constant but regulated speeds can be obtained by an adjustment of the pointer on P. For large motors electronic regulation will require an amplification of the output of the

regulator tube using one of the methods covered in the preceding discussion of voltage regulation. A complete system of d-c motor-starting operation and regulation by electronic means was treated in Chapter XVI.

Temperature Regulation. The maintenance of a constant temperature is often necessary in continuous-flow processes. Such temperature regulation is achieved by the principle of servomechanisms. When

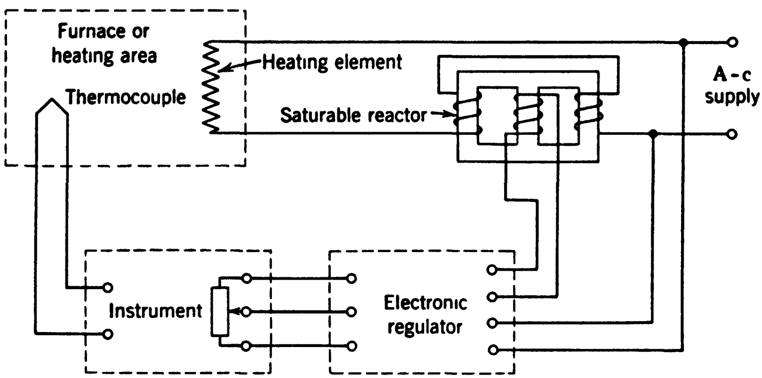


FIG. 8. Block diagram for temperature regulation.

electrical energy is employed for supplying the heat, a circuit following the block diagram of Fig. 8 is generally employed. The temperature-detecting device may be a thermocouple, a phototube (for high temperatures) or a resistance unit which has a high temperature coefficient of resistance. The instrument is usually a device that indicates the temperature and, in addition, has the ability to transfer the temperature-indicating signal into a potential difference for actuating the electronic regulator. The electronic regulator is a thyratron and has a free-wheeling circuit for energizing the d-c winding of the saturable reactor. The saturable reactor controls the current to the electric heating element.

REFERENCE

BENDZ, W. I., and C. A. SCARLOTT, *Electronics for Industry*, John Wiley & Sons, 1947.

Chapter XIX

ELECTRONIC POWER CONTROLS

One of the most useful methods of controlling electric power utilizes the saturable reactor in conjunction with electronic tubes and circuits. Some applications of this method constitute regulation wherein some factor such as voltage, speed, or temperature, may be held constant. In other cases the control is applied manually or serves as a limiting factor. Three of the applications that follow utilize a saturable reactor placed in series with the supply line as the direct limiting device for controlling the load.

Battery-Charger Control. Batteries charged by electronic rectifiers will become overcharged unless some provision is made for limiting the degree of charge. One simple method of providing such limitation is shown in the circuit of Fig. 1. Here a saturable reactor *SR* is connected in series with the primary of a transformer *T*. The secondary of the transformer with a mid-tap supplies voltage to two phanotrons for full-wave rectification and charging of the battery *B*. The voltage-limiting detector consists of a magnetic solenoid type of relay *S* connected in series with an adjustable resistor *P* across the battery terminals. The solenoid relay controls make and break contacts (held closed by helical spring action) in the supply to the d-c winding of the saturable reactor. When the battery voltage is low, the contacts are closed, the saturable reactor offers little impedance to current flow, and the battery is being charged. When the battery voltage rises to the desired value the solenoid opens its contacts and charging ceases. The terminal voltage of the battery at which the charge is cut off may be adjusted by the dial for resistor *P* or by a change in spring tension.

Electronic Furnace Control. The temperature of an electrically heated furnace may be limited for each new charge added by a saturable reactor and electronic control circuits. In a continuous-flow process, the temperature may be regulated as shown in the block diagram of Fig. 8 in Chapter XVIII. A schematic circuit that may be applied to both types of application is given in Fig. 2. A saturable reactor

SR is placed directly in series with the furnace heater. The d-c winding of the saturable reactor is excited by thyratron 1 and phanotron 2 connected in a free-wheeling circuit to assure a continuous direct current (see page 228 for free-wheeling circuit). The firing of thyratron 1 is controlled by a d-c grid bias provided by the combined diode-triode tube 3. Part of this bias is provided by the rectifier action of

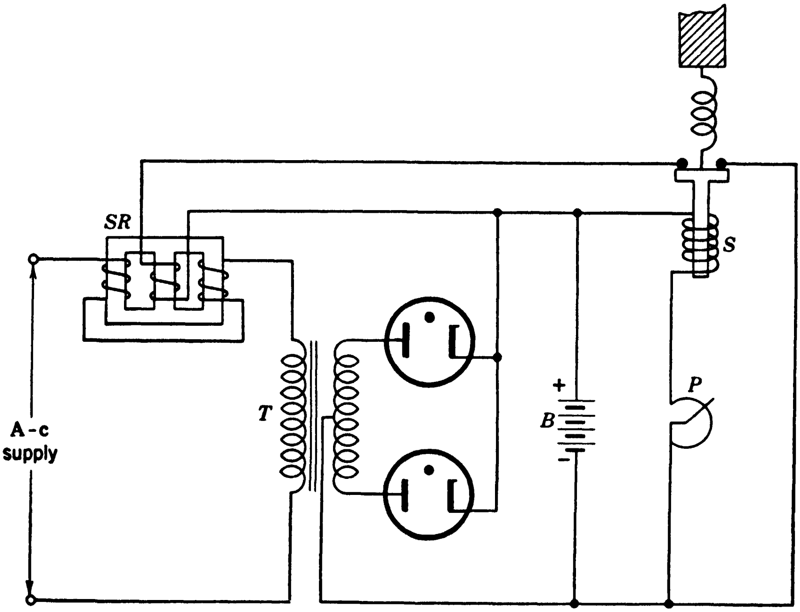


FIG. 1. Circuit for charging a storage battery.

the diode (left side of tube 3), giving an electron flow and polarity as shown for resistor *R*1. The other part of the bias is provided by the rectified current of the triode (right) in passing through resistor *R*2. The magnitude of the latter current is controlled by the potential applied between the grid and cathode of triode 3. This potential is provided by transformer *T*1 which lies in the balance or middle arm of the bridge circuit at the right. In this bridge circuit, the resistances of arms *A* and *B* are located on an instrument and their magnitudes are controlled by some heat-indicating device within the furnace. As *A* and *B* change, the bridge is unbalanced and the voltage across *T*1*P* (primary) induces a voltage across *T*1*S* (secondary) which controls the grid of triode 3. From the preceding statements, it is apparent that the potential drops across *R*1 and *R*2 are in opposition and that

their difference is applied to the grid of thyratron 1. The voltage drop across *R*1 is produced by transformer *T*2. The excitation of *T*2 is almost entirely from the upper potentiometer which makes it a negative or degenerative feedback. This tends to improve the linearity of the relationship between the triode output and the lead voltage. A slight amount of voltage is obtained from the lower potentiometer so that at

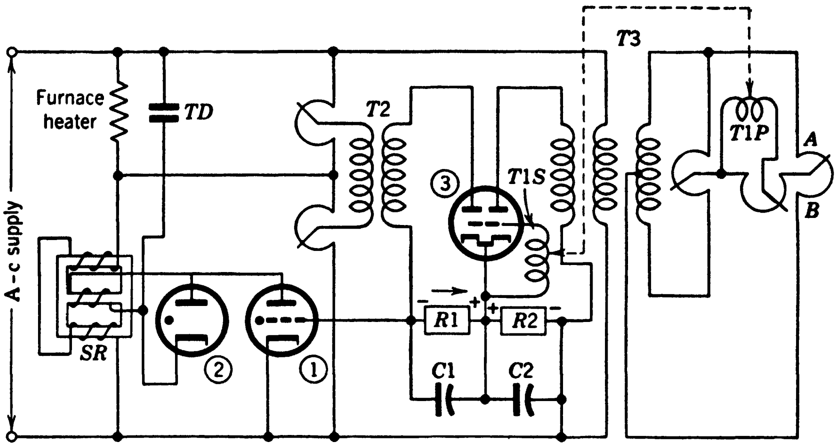


Fig. 2. Temperature regulation of an electric furnace.

very light loads sufficient voltage exists across *R*1 to turn the thyratron completely off when the *R*2 voltage goes to zero.

Theater Light-Dimming Control.* The stage of a modern theater requires a large number of separate groups of lamps for producing the desired lighting effects. Such groups may be near the front of the stage (footlights), or above, or at the side. In each location, groups of different color are required and in order to produce the proper gradations of color the illumination furnished by each group must be varied in a stepless manner. The amount of light given forth by an incandescent lamp varies with the applied voltage. Thus incandescent lamps may be dimmed by placing a rheostat in series. This method produces heat within the rheostat and wastes power. With many groups of lights and rheostats the amount of heat to be dissipated may present a problem. Hence a superior method for controlling the groups of lamps for stage lighting is to employ a saturable reactor for each group and to control its action by electronic circuits. The transformers, re-

* For a complete treatment of this subject see E. D. Schneider, "Thyratron Reactor Lighting Control," *Elec. Eng.*, June 1938.

actors, and most other parts of the circuit may be placed in any convenient location beneath the stage or elsewhere. The manual-control mechanism consists of a small dial or lever type of resistor which may be remote from all other equipment. Thus it is possible to place all the controls for a large theater on a relatively small panel which may be located in the orchestra pit or even in the balcony. Small lever-type control resistors may be ganged together to give various combinations of group control.

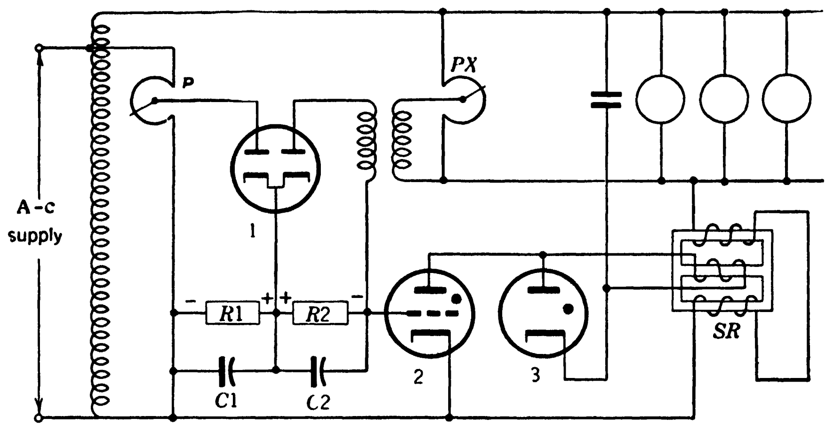


FIG. 3 Circuit for a theater lighting control.

The circuit for the control of a single group of lamps is given in Fig. 3. This circuit is similar to that of Fig. 2 in many respects, differing mainly in its manual form of control. Power is fed from an auto transformer on the left to the saturable reactor and lamp load on the right. The auto transformer steps up the voltage to compensate for the voltage drop in the reactor and to give full rated voltage across the lamps. The d-c winding of the saturable reactor is fed by a free-wheeling circuit. The firing of thyratron 2 is controlled by a d-c bias voltage applied to its grid. This bias voltage is the difference in voltage drops across resistors $R1$ and $R2$. The drop across $R2$ is determined by the lamp voltage while that across $R1$ is controlled by the setting of the arm on potentiometer P. With potentiometer P set full on (up), PX is adjusted to give the normal voltage of the lamps. As the arm of P is moved down, the voltage drop across $R1$ lowers, the thyratron grid becomes more negative, tube 3 fires later in the cycle, less current passes through the d-c winding of SR, and the voltage across the lamps falls. As the lamp voltage falls, less current is recti-

fied by right diode tube 1, which gives a reduced drop across $R2$ and permits a new balance of circuit conditions. The placing of PX across the lamp group (lamp voltage) instead of across the a-c line assures, the same lamp voltage for a given setting of P regardless of the number of lamps in a group. This permits the use of the same equipment for lamp groupings having different loads in watts.

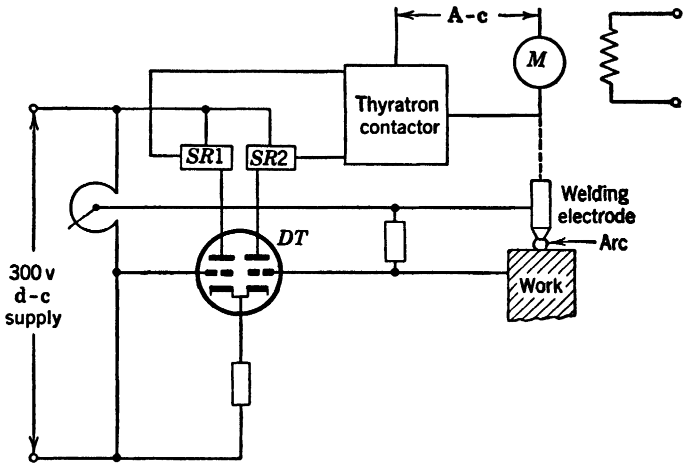


FIG. 4. Schematic diagram of an arc-welding control.

A smooth phase control of the thyratron is effected in Fig. 3 by adjustment of the time constant of $R1C1$ so that an appreciable ripple exists. With correct phase relations and adjustment this ripple will intercept the critical grid-voltage curve (firing) so as to give a smooth control from "full on" to "full off." (See Fig. 13, Chapter VI.)

Arc-Welding Control. Most arc welding is performed by an operator who maintains a suitable electric arc between a metal rod or electrode and the work being welded. A good operator maintains the proper length of arc and moves the arc along the work at the proper rate. For long line welds and for repetitive welding jobs an automatic welding system may be designed to reduce the labor of the operator, produce more uniform results, and reduce the cost. An automatic system utilizes two electric motors, one for maintaining the proper length of arc and one for feeding the arc along the work. A good weld requires the proper current flow in the arc and the right amount of energy delivered into the arc. Since both these factors vary with the voltage drop across the arc, the simplest method of automatic arc control lies in governing the potential drop across the arc. The arc drop

in turn depends on the distance between the electrode and the work.

A schematic diagram illustrating the general principle for controlling the welding electrode is given in Fig. 4. The distance between the electrode and the work (right) is regulated by a reversible d-c motor *M*. This motor has a constant field produced by rectified alternating current. The motor armature may receive rectified power from either of two thyratrons, each of which tends to drive the motor in an opposite direction. The thyratron contactor unit is controlled by electronic circuits designed to maintain the proper arc drop. In these circuits the voltage drop across the arc controls the grid potential of one unit of

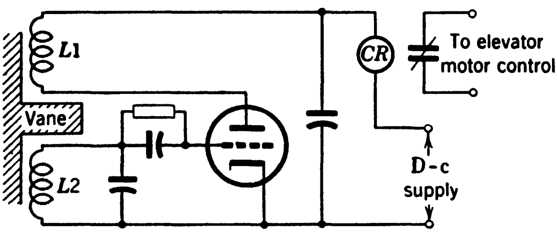


FIG. 5. Oscillator circuit for elevator-car leveling.

a double triode, while the grid potential of the other triode is determined on a potential divider across a constant d-c circuit. The basic control circuit is the long-tailed pair described on page 239. The cathode-anode currents of the triodes supply the d-c winding of two saturable reactors *SR*1 and *SR*2 which, in turn, provide the phase shift for firing the thyratrons. The output of the thyratrons acts in opposition to supply rectified d-c voltage to the armature of the motor. For a proper value of arc drop and length, the opposition forces are balanced and the motor is at rest. As the arc drop varies from normal an unbalance is set up in the servomechanism circuits which causes the motor to operate in a suitable direction to restore the arc drop to normal.

Elevator-Leveling Control. The elevators in modern buildings are usually fully automatic, leaving only the function of floor selection via push button to the operator. One of the steps in automatic elevator operation is the proper leveling of the car as it stops at the selected floor.

One method of automatic floor leveling uses a simple oscillator relay unit with a circuit as shown in Fig. 5. This is a tickler oscillator circuit wherein the necessary feedback is provided by the transformer action between coils *L*1 and *L*2. With the proper selection of com-

ponent constants and in the absence of the vane shown in the figure, the circuit oscillates continuously. While it is in oscillation, the reactance arising from the a-c frequency limits the current through the relay *CR* so that it does not energize and its contact remains closed. Whenever the magnetic vane is interposed between coils *L*1 and *L*2, the coupling is broken, feedback stops, and the circuit ceases to oscillate. When oscillation stops, direct current flows through the cathode-anode of the triode and the winding of relay *CR*. Since *CR* offers little opposition to direct current, the current rises quickly to a value that energizes relay *CR* and opens the contact to the elevator motor control.

FIG. 6 Automatic elevator-car-leveling apparatus with photron units mounted on the elevator car and vanes mounted in the hatchway.

In the application of this oscillator relay control several units employing the circuits of Fig. 5 are mounted on top of the elevator car. These units are mounted so that as the car moves vertically the space between coils *L*1 and *L*2 is intercepted by fixed vanes mounted along the walls of the elevator shaft. A suitable vertical location of these vanes will control the operation of the oscillator relays in steps so that the resulting signals to the control circuits for the elevator drive will decelerate the elevator car at the proper rate and bring the car to rest at the exact floor level. An illustration of the car-leveling oscillator unit is given in Fig. 6.

REFERENCES

CHUTE, G. M., *Electronics in Industry,* McGraw-Hill Book Co., 1946.

BENDZ, W. I., and C. A. SCARLOTT, *Electronics for Industry,* John Wiley & Sons, 1941.

Chapter XX

AMPLIDYNE SERVOMECHANISMS

Most of the applications of control treated in the preceding chapters have been effected through the use of electron tubes and circuits. The applications of Chapter XIX employed the saturable reactor as the important link in control. The applications of this chapter will feature the use of the rotary amplifier as the basic unit for power control. The rotary amplifier is frequently used in servomechanism systems for controlling larger units of electrical power and in systems where large-capacity rotating machines are employed. Any one of the forms of rotary amplifier units discussed in the last section of Chapter X may be used in control applications but only one form will be depicted in the applications that follow. Rotary amplifiers may be employed to control voltage, current, speed, power factor, position, or combinations of one or more of these factors simultaneously. Some of the practical combinations are voltage control with current limit, current control with voltage limit, speed control with voltage and current limit, and positioning with voltage and current limit.

Voltage Control. The amplidyne may be employed to control the voltage of an a-c generator by serving as the exciter for the generator field. A schematic circuit for this application is shown in Fig. 1. Here the a-c voltage is rectified by a bridge-type copper oxide rectifier circuit and impressed across a potential-dividing resistor. Another source of a-c supply is rectified and loaded by a resistor (lower part of Fig. 1). The voltage drops across the two resistors are connected so as to feed in opposition the control field and the standard reference field of the amplidyne. Any voltage difference between these *RI* drops causes a change in flux in the magnetic field of the amplidyne. This change of flux is in the proper direction to control the amplified voltage output of the amplidyne so that the alternator voltage will be restored to the desired normal value. The output voltage of the alternator may be set for different values by means of the voltage adjustment on the potential divider.

Other applications of voltage control using rotary amplifiers may involve quick changes of voltage or even reversal of voltage in Ward Leonard motor drives in steel mills and the mining industry.

Current Control. The current in a motor or generator circuit can be held within close limits by the amplidyne regardless of speed, voltage, or load changes. Current control may be desired (1) so that the machine may be operated at peak performance without danger of over-

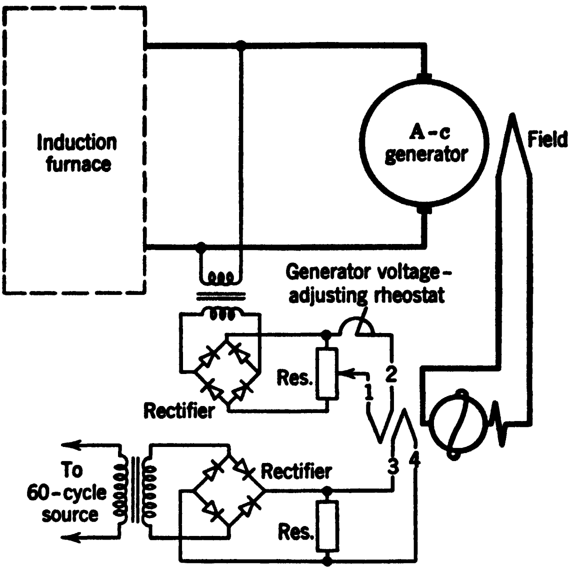


FIG. 1. Circuit for an amplidyne used with a high-frequency generator for induction heating. (Courtesy General Electric Company.)

load, (2) so that maximum rates of motor acceleration or deceleration may be obtained, or (3) in order to provide a constant tension upon a wire, steel strip, or other material while the motor is in continuous motion. One application falling under class 3 uses the circuit shown in Fig. 2. This diagram illustrates how a large reel motor can be made to operate over a wide range of speed, automatically holding constant the current and tension on the reel. In this application the winding reel for providing tension is driven by a reel motor which has an armature supply from a separate reel generator and a field supply controlled by an amplidyne. The torque and tension are controlled by varying the field strength of the reel motor. The controlling amplidynes employ four separate fields. Field 3–4 is a reference field

which tends to increase the torque of the reel motor. Field 5–6 is the control field for the large exciter-amplidyne and is fed by a small control amplidyne. Field 2 of the latter amplidyne is fed by the *IR* drop in the reel-motor commutating field. Fields 3–4 and 5–6 are in opposition and a balance is quickly reached. A limit field (7–8) is

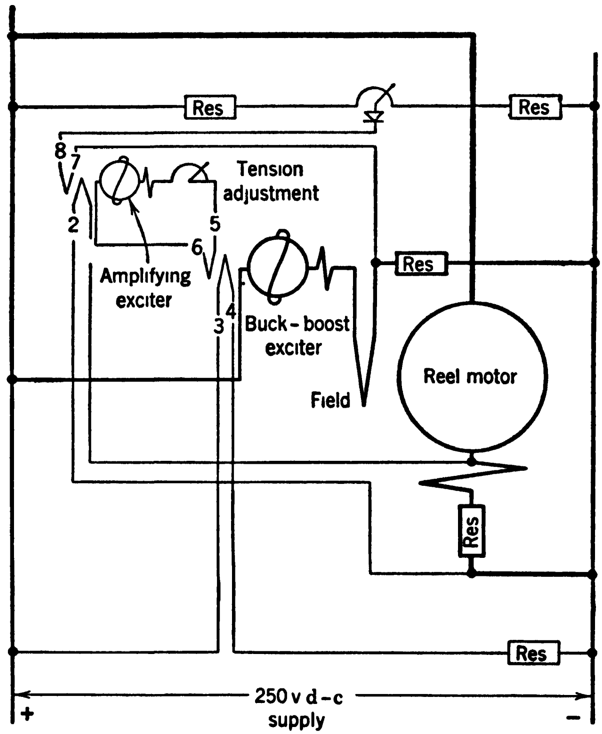


FIG. 2. Circuit for amplidyne control of a reel motor on a cold-strip mill (Courtesy General Electric Company)

normally inactive. Its function is to strengthen the reel-motor field whenever the reel-motor current drops below a minimum value.

Voltage Control with Current Limit. Amplidyne control can be adapted to applications such as power shovels, mine hoists, ship hoists, saw-mill carriages, and similar loads. For these applications it is desirable to apply maximum safe current to the motors during acceleration, deceleration, and operation of loads and to limit the voltage when the loads become light or overhauling in character. A typical circuit for this class of duty is given in Fig. 3. The power for driving

the motor or two motors in some applications is supplied by a separate d-c generator. The field for this generator is excited by an amplidyne. The amplidyne has three fields for providing the desired controls. The first field is a constant reference field 3–4 excited from a d-c supply shown at the bottom of Fig. 3. A second field 5–6 for providing cur-

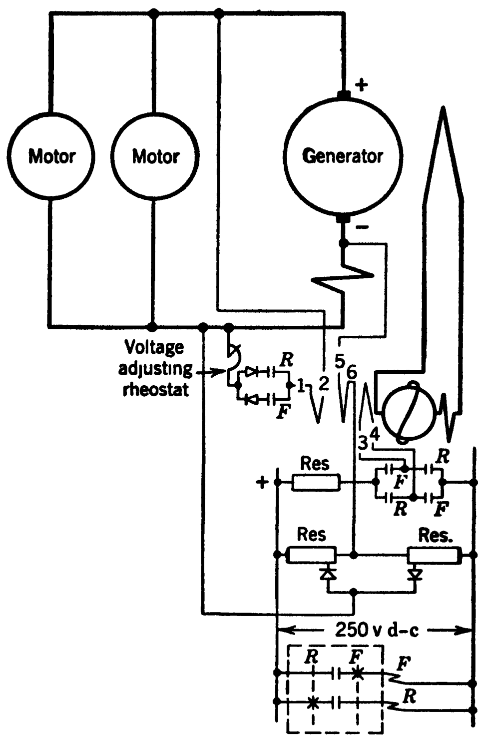


FIG 3. Circuit for an amplidyne system providing voltage control with current limit. (Courtesy General Electric Company)

rent control is excited by two voltages in opposition. One of the voltages is proportional to the load current and is produced by an *RI* drop across the resistor in series with the generator armature. The second voltage is provided by the drop across a potential divider energized by a rectified a-c voltage. The difference voltage in this series current control circuit limits the current output of the generator. A small rectifier in the current control field prevents a current flow until the voltage across the armature resistor exceeds that produced by the rectified supply. The third control field 1–2 for the amplidyne is connected across the generator terminals and carries a current in such a

direction as to limit the voltage rise of the generator under light or reverse loads.

In the use of this system the motors are reversed by reversing the polarity of the generator. This change of polarity requires a reversal of some connections in the reference field and the voltage control field. These changes are made by relay contacts marked F and R in Fig. 3

FIG. 4. Electric power shovel equipped with an amplidyne control. (Courtesy General Electric Company.)

for forward and reverse direction of rotation. A power shovel using this type of control is illustrated in Fig. 4.

Positioning Control. Industrial processes often require the control of the position of mechanical units for efficient and satisfactory operation. The control of the length of the electric arc and the leveling of the elevator car cited in Chapter XIX are good examples of positioning control. An example employing the amplidyne is the automatic adjustment of the electrodes in an arc furnace. An arc furnace employing such control is illustrated in Fig. 5 and the schematic circuit for effecting such control is given in Fig. 6. Each of the graphite electrodes of the three-phase arc require the circuit and components shown in Fig. 6. The electrodes are raised and lowered by the electrode motor shown at the top and this motor is powered directly by an ampli-

dyne. The main control field of the amplidyne is excited by two voltages in opposition. One of these voltages is the rectified a-c arc drop and the other is the rectified voltage from the secondary of a current transformer. Since the power delivered to the arc is proportional to

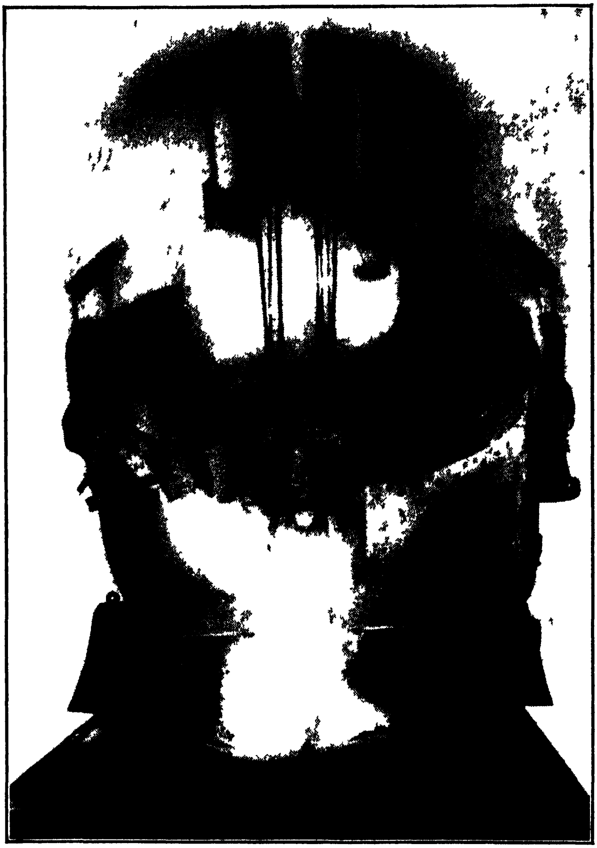


Fig. 5. Metal-melting arc furnace with an amplidyne control. (Courtesy General Electric Company)

the product of the phase current and the arc drop, it is obvious that the proper balance between these two factors must be maintained for best results. The desired load may be secured through the setting of the load adjustment potentiometer. To provide manual adjustment of electrodes when a new charge is being placed in the furnace, a hand control reversing switch supplies power to the hand control field of the amplidyne. When the furnace is turned on after a fresh charge has been added, there are likely to be transient fluctuations of arc voltage

and current which would keep the control mechanism hunting in an unnecessary manner. Such hunting is reduced by an antihunt control field (not shown) which acts in opposition to the main control field.

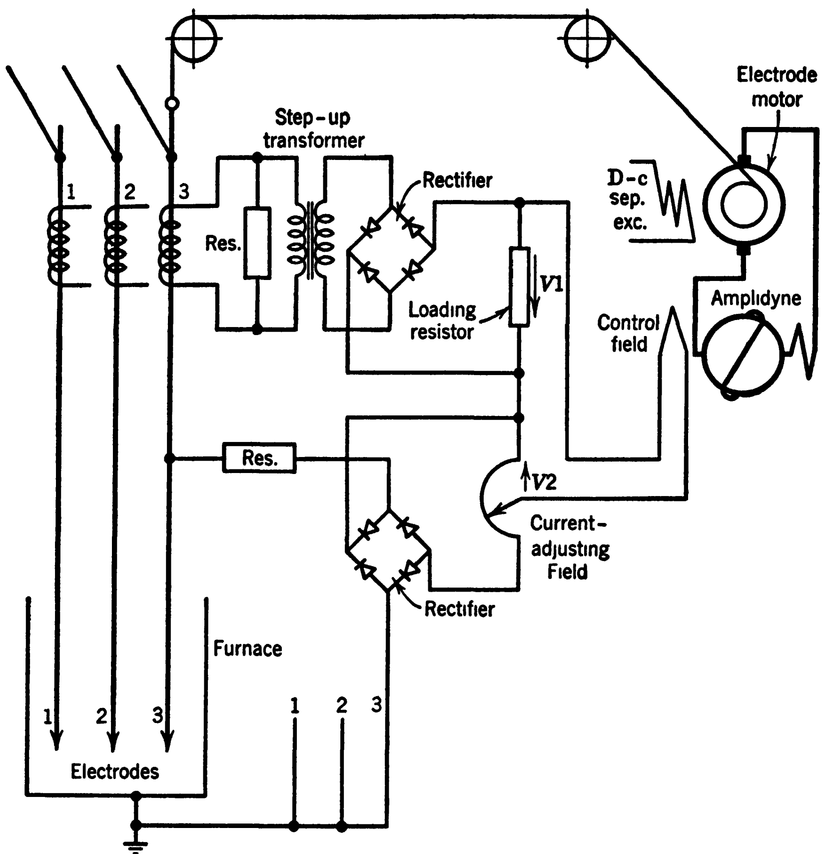


FIG. 6. Circuit showing power control by positioning obtained with an amplidyne on an arc furnace. (Courtesy General Electric Company.)

Speed Control. The amplidyne can be the electric coupling that matches the speeds of two or more d-c motors, or it can provide a simple means of holding speeds to desired values. In these applications speeds are translated into voltages by tachometer generators, and these voltages are applied to the amplidyne in the same manner as shown in Fig. 1. With this arrangement, (1) speed can be changed rapidly, (2) low speeds are held accurately, and (3) accuracy of speed regulation is assured.

One important use of amplidyne speed control is applied to the flying shear for cutting the sheet-steel strip into uniform lengths after it emerges from the rolls at speeds up to 2000 feet per minute. A schematic circuit for this application is shown in Fig. 7. In this circuit the difference voltage between the tachometers of the standard motor and the shear motor is amplified by an electronic regulator to excite one

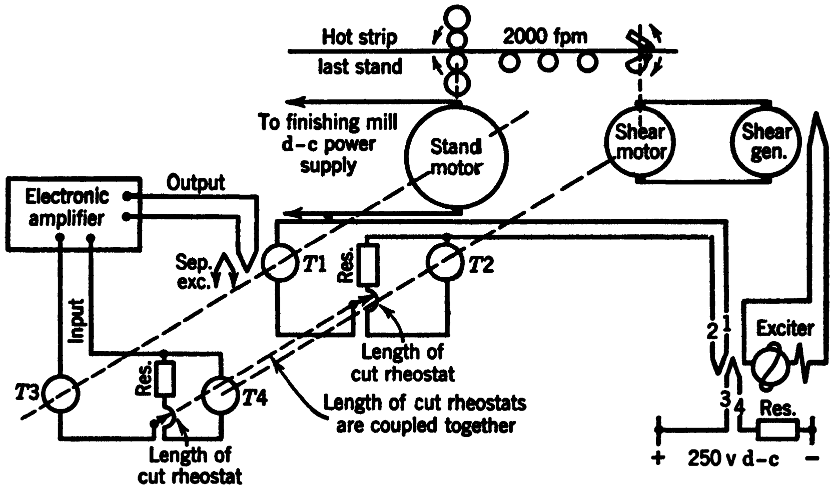


FIG. 7. Circuit for an amplidyne-controlled flying-shear speed control. (Courtesy General Electric Company.)

field of a small amplidyne tachometer. The other field of this amplidyne is the usual constant reference field. The output of this amplidyne is balanced against the voltage of the generator in series with the field of a second larger amplidyne exciter which controls the field of the generator. The remaining explanation of this circuit follows the theory of the preceding examples. In passing it is well to note that there is a trend toward using electronic excitation of amplidyne fields in order to improve accuracy, attain stability, and increase the speed of response in the more critical applications.

REFERENCE

ALEXANDERSON, E. F. W., M. A. EDWARDS, and K. K. BOWMAN, "The Amplidyne Generator—A Dynamoelectric Amplifier for Power Control," *Gen. Elec. Rev.*, March 1940.

Chapter XXI

X-RAY APPLICATIONS

X-Ray Tube. The X-ray tube is a two-electrode tube invented by Roentgen in 1895. This tube had the ability of producing electromagnetic radiations known as roentgen rays. In 1913 Doctor Coolidge made improvements which have resulted in wide and valuable applications of the tube. All modern X-ray tubes use a heated tungsten cathode partially surrounded by a cylinder or focusing cap for directing the electron emission toward the anode. The anode, generally called the target, is made of tungsten or of a tungsten insert in a copper base. The two electrodes are encased in a highly evacuated Pyrex glass tube as illustrated schematically in Fig. 1. The face of the target is at an angle to the axis of the tube so as to direct the emanating rays in a direction where they may be utilized.

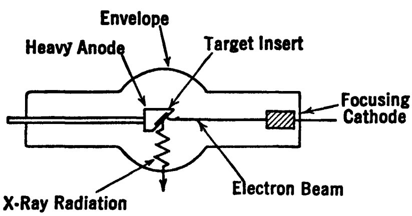


FIG. 1. Construction of an X-ray tube.

The electrons emitted by the cathode are attracted to the anode by a high potential of the order of 10,000 to 1,000,000 volts and sometimes higher. The tremendous impact of the electrons in hitting the target excites the electrons in the atoms of the target so that they send out electromagnetic radiations known as X rays. These emanations are not electrons or particles but wave energy similar to light waves. It will be recalled that the impact of electrons on atoms of gas produces an excitation of the electrons which results in the production of light. In the X-ray tube the impact of the electron is much greater because of the high potential gradients and it acts upon electrons in a solid so that the resulting waves are much shorter (higher frequency) and contain a greater energy. It is probable that the impinging electrons in the X-ray tube penetrate into the inner electron rings and even the

nuclei of the atoms constituting the target. The position of the X rays in the electromagnetic spectrum was shown on page 159. The X rays emitted by the tungsten target are a continuous spectrum of electromagnetic radiation. The minimum wavelength is determined by the peak voltage across the X-ray tube according to the equation $Vl = 12.354$, where V is the voltage across the tube measured in kilovolts and l is the minimum wavelength in angstrom units (10^{-8} centimeter). X rays are propagated with the same speed as light, follow the inverse square law, are unaffected by electric and magnetic fields, and can be reflected, diffracted, and polarized.

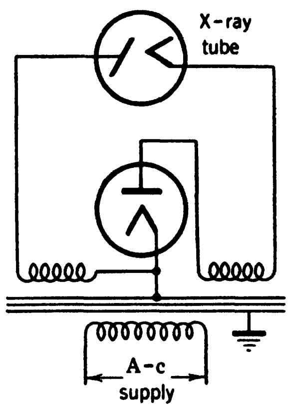


FIG. 2. Half-way rectifier for an X-ray tube.

X rays have the power of penetration through layers of solids that are opaque to light. These X rays are invisible to the eye but can be detected after passing through solids by their action on a fluorescent screen or on a photographic film. Their ability to blacken sensitized film in a manner proportional to their intensity has been the basis of radiography. Furthermore, their ability to penetrate opaque objects is dependent upon the atomic weight (density) of the object. The penetrating power of X rays is referred to as a degree of hardness; soft rays have small penetrating power, hard rays deep penetrating power. Hardness is proportional to the frequency of radiation and depends on the voltage applied across the cathode-anode circuit. The penetrating power varies approximately as the square of the voltage. Thus in the operation of X-ray devices the hardness is controlled by varying the cathode-anode voltage, the intensity or quantity of X rays is controlled by the voltage across or the current passed through the cathode for heating, and the quantitative effect of the X rays is controlled by the time of exposure.

X-Ray Equipment. An X-ray machine consists of a high-voltage power supply, an X-ray tube, and control equipment. Since X-ray tubes are rectifiers, they may be energized by alternating current, by half-wave or full-wave a-c rectifiers, or by direct current. Since a heavy load may raise the temperature of the tungsten target to a point where appreciable electron emission may occur, there are definite limits on the use of half-wave and a-c excitation if inverse current

flow and tube destruction are to be avoided. A conventional circuit for half-wave rectification is shown in Fig. 2. It should be noted that the kenotron rectifying tube and the mid-point of the high-voltage secondary of the transformer are grounded. A second circuit for half-wave rectification is given in Fig. 3. The auto transformer on the left provides for a control of the magnitude of the peak applied voltage.

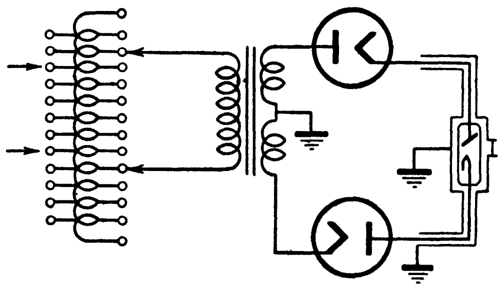


FIG. 3. Two-tube half-way rectifier with a shockproof X-ray tube.

The grounded mid-point of the high-voltage transformer winding limits the voltage to ground at any point in the circuit to one-half of the a-c peak. The use of two kenotrons reduces the inverse peak per tube and prevents any inverse electron flow through the X-ray tube. Full-wave rectification for the X-ray tube may be secured by the use of a transformer with mid-tap and two kenotron rectifier tubes or by the use of the bridge circuit with four kenotrons as covered in Chapter XII on power rectification and inversion. The maximum cathode-anode voltage across the X-ray tube for the preceding circuits for power supply is the a-c secondary high-voltage peak. A special power-supply circuit which applies two times the peak voltage across the X-ray tube is shown in Fig. 4. This is known as the Villard circuit and the explanation of its operation is left to the reader.

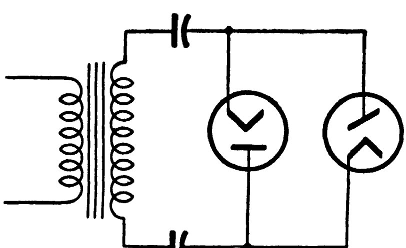


FIG. 4. Villard circuit for X-ray equipment.

Since X-ray tubes are operated at the saturation current, the electron current through the tube is determined by the size and temperature of the tungsten filament. The tungsten anode or target may be mounted on the end of a molybdenum rod or set in a copper anode.

Since only a fraction of one per cent of the electrical power put into the X-ray tube is converted into X rays, nearly all the remainder appears as heat in the target. Therefore, effective means of cooling the target must be provided. Those anodes that consist only of tungsten mounted on molybdenum rods are cooled by radiation. Those targets that are cast in copper rods are cooled by heat conduction along the copper rod to a radiator which permits the heat to be carried away by air or by oil. The anodes of some tubes which are operated continuously at high loads are hollowed out behind the target and cooled by circulating water or oil. The minimum length of an X-ray tube is

Fig. 5. X-ray photographs of human teeth: *left,* impacted tooth and fillings; *center,* abscessed tooth; *right,* permanent teeth pushing up baby teeth. (Courtesy Dr. C. J. Buster)

determined by the spark-over distance for the maximum voltage applied to the tube. Many modern X-ray tubes are made shockproof by being operated inside grounded metal tanks filled with oil. These tubes are shorter than those operated in the open air. A shockproof tube is indicated in Fig. 3.

The controls for X-ray machines govern the voltage applied at the X-ray tube and the time the power is applied. These controls follow the general principles covered in the preceding sections of this text.

Applications of X Rays in Medicine. X rays are used in the field of medicine for radiography and therapy. From a knowledge of anatomy and the relative opacities of the various organs of the human body to X rays, it is possible to recognize injuries or diseases by examining the radiograph (shadows upon a photographic film). Broken or dislocated bones, defective teeth, and the presence of bullets or other foreign material are shown by the radiograph. Examples of radiographs of teeth are given in Fig. 5. Lung tissue destroyed by tuberculosis, gallstones, and ulcers of the stomach may be detected on a radiograph by the radiographic specialist. There is an optimum voltage for radiographing each part of the human body depending on

the thickness and density of the part viewed. The maximum voltage required for radiographing the human body is 100,000 volts. X rays are used to treat or destroy malignant tumors such as cancer. For treating cancers near the surface, tube voltages varying from 10,000 to 140,000 volts are satisfactory, but for deep-seated tumors very hard X rays produced by tubes operating at 200,000 to 1,000,000 volts are required. Since these high voltages also produce soft rays which would burn the surface tissue, these softer rays are filtered out by sheets of copper and aluminum. Radiographic tubes are operated intermittently

FIG. 6. X-ray tube for deep therapy. (Courtesy Westinghouse Electric Corporation.)

for comparatively short periods of time, whereas tubes for therapy must be capable of carrying currents of 10 to 25 milliamperes continuously. A deep-therapy X-ray tube is illustrated in Fig. 6.

Industrial Applications of X Rays. In industry X rays are used for fluoroscopy, X-ray diffraction, and radiography of machine parts. *Fluoroscopy* is the visual representation of the construction of opaque objects on a fluorescent screen by the transmission of X rays. This process has wide application in the inspection of foods and packaged devices before distribution. Under small-scale methods the parts to be inspected may be placed under the fluorescent screen manually whereas under large-scale processes the parts to be inspected move on a conveyer belt so that an operator makes a visual inspection as the shadowed view passes on the fluorescent screen. A fluoroscopic inspection machine with conveyer belt is shown in Fig. 7. Operators stationed on either side watch the lines of oranges passing under the screen and eliminate bad ones by operating levers which eject them from the line. A fluoroscopic picture of a good and a defective orange is shown in Fig. 8. Apparatus of the type described is widely used for inspecting all kinds of citrus fruit, and in canning plants for detecting damaged vegetables and foreign bodies, and for checking the fill of packaged containers. Tube voltages somewhat under 100,000 volts are used in fluoroscopy.

The nature and behavior of most substances depends upon the arrangement of atoms and molecules in the crystalline structure. This

FIG. 7. Fluoroscopic inspection device with a conveyer belt. (Courtesy General Electric X-Ray Corporation.)

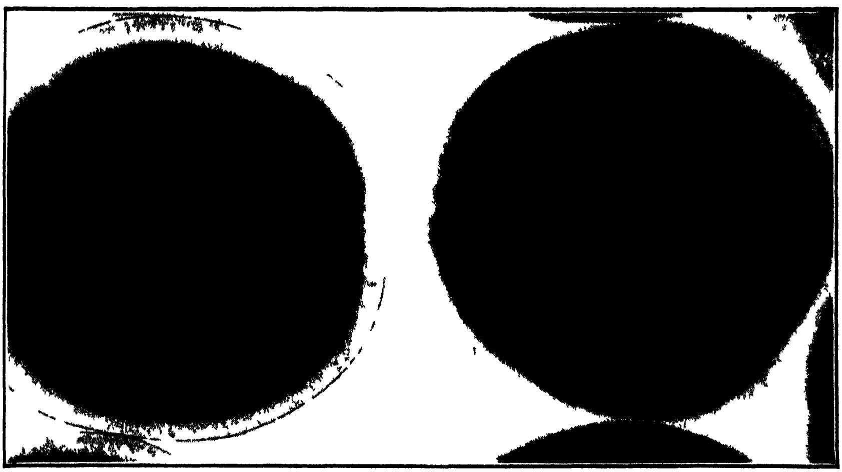


FIG. 8. X-ray photograph of oranges: *left,* good orange; *right,* defective orange. (Courtesy General Electric X-Ray Corporation.)

is true whether the material is a deep drawing steel, a bearing metal, or a tempered steel spring. Every substance such as rolled, forged, or heat-treated metal has a distinct atomic arrangement which determines its properties. Each atomic arrangement controls the diffractive effect produced by X rays. Thus, if a beam of X rays is passed through a crystal, the beam will be bent or redirected in a series of emergent rays whose separation and intensities are characteristic of the material. A radiograph of the diffracted rays constitutes a "fingerprint" of the substance because no two substances have been found to produce identical diffraction patterns. *X-ray diffraction* is used (1) as a laboratory tool in the hands of a trained specialist working on problems of production or product research, and (2) as a routine inspection tool in the hands of an operator making repetitive checks on the analysis or condition of a material. Three methods are employed in making analyses by X-ray diffraction. In one method X rays are passed through a very thin section of the substance and the diffraction pattern falls on a vertical film. In a second method the X rays are projected upon a powdered sample at the center of a drum and the film is placed inside the drum. The third application, called a back reflection method, determines the *reflection* pattern from the test specimen. The back reflection method is employed with alloys and combinations of materials that form a substantial solid solution. Examples of X-ray diffraction patterns are given in Fig. 9. Voltages used on X-ray tubes for diffraction studies vary from 10,000 to 50,000 volts. An X-ray diffraction tube is shown in Fig. 10.

An important application of X-ray diffraction methods is the determination of the optimum angle for cutting quartz to produce wafers (crystals) for controlling the frequency of crystal oscillators. The orientation or angle requirement for crystals manufactured before the war was met by a trial-and-error performance-selection method. During the war more crystals were required per day of manufacture than in an entire year before. Also it was necessary to produce crystals that would give the same frequency regardless of temperature.

These requirements were met by the development of an X-ray goniometer which indicates the intensity of X-ray diffraction in the direction characteristic of a certain set of planes. A sample crystal wafer is turned in the X-ray beam until this set of planes diffracts with maximum intensity. At this position the orientation of the diffracting planes is determined and the proper angle for cutting the mother quartz is determined.

Closely allied to X-ray diffraction methods is the process of microradiography. Here radiographs are made of exceedingly thin sections of materials. These radiographs can be greatly enlarged to provide a means of adding depth to the studies of surface conditions with which

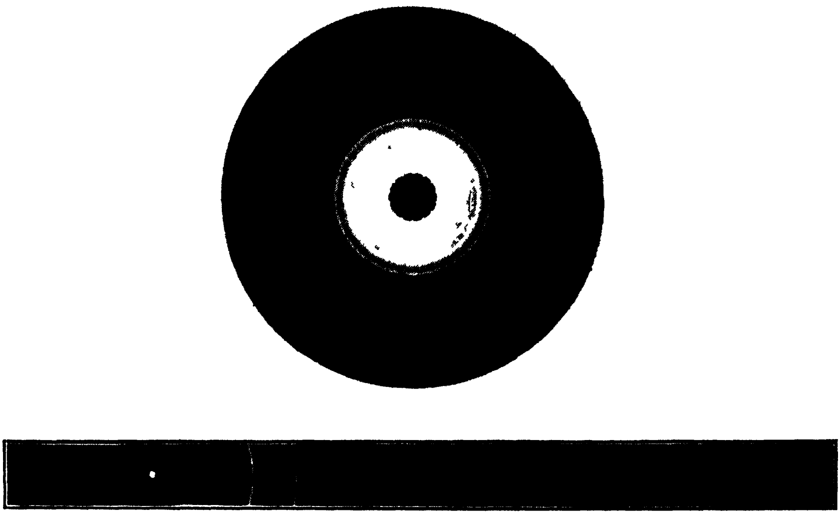


FIG. 9. X-ray diffraction pattern of salt containing 96.12 per cent potassium chloride (*top*) and cold-worked stainless steel. Outer dot shows anstenite line; inner dot, ferrite (*bottom*). (Courtesy General Electric X-Ray Corporation.)

microphotography is concerned. Radiographs can be applied to solutions so that the process has applications in chemistry as well as in metallurgy. These applications require tube voltages of the order of 6000 to 9000 volts.

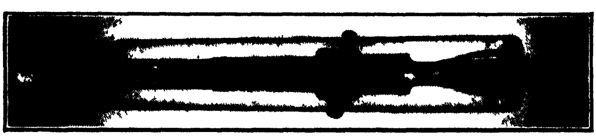


FIG. 10. X-ray diffraction tube. (Courtesy General Electric Company.)

X-ray inspection with radiography of castings, machine parts, welding structure, and intricate machinery assemblies provides the design engineer and the production department with a tool of incalculable value. Through the use of X-ray inspection, casting and electric welding techniques may be determined for producing stronger and flawless

castings and electric welds. Inspection of castings before machining permits the elimination of parts that would result in inferior or rejected products. X-ray inspection permits a complete study of mechanical parts that are ordinarily obscured. It is important to observe that this method is a nondestructive inspection.

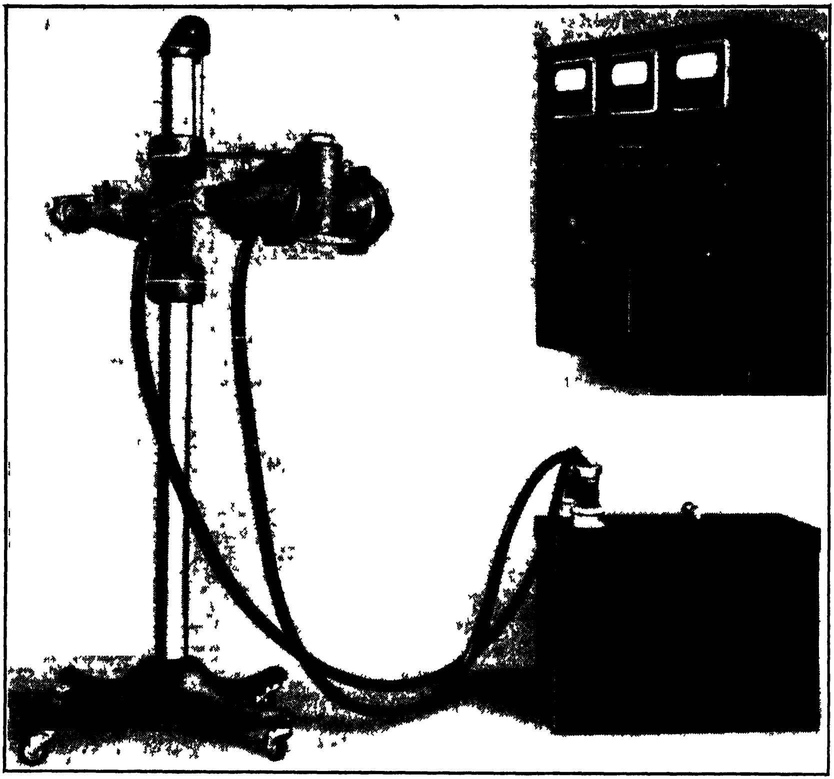


FIG 11 Industrial X-ray unit and control, 150 kilovolts (Courtesy Westinghouse Electric Corporation)

X-ray radiographic inspection utilizes a wide range of applied tube voltage and equipment. For aluminum castings and iron or steel parts up to ½ inch in thickness, a range of 60 to 140 kilovolts may be satisfactory. A 150-kilovolt industrial X-ray unit is illustrated in Fig 11. The X-ray tube in this unit is of the shockproof, rayproof, oil-immersed design provided with a water-cooling coil. The tube housing is made of cast aluminum and has a bellows arrangement to accommodate all normal expansion of the insulating oil. The expansion bellows are provided with a warning indicator which signifies the approach of

unsafe operating temperatures. The interior of the tube housing contains a copper cooling coil which is so terminated that connections to a tap water supply may be made. The X-ray tube head is arranged so that it may be rotated through an angle of 270 degrees around its long axis to give radiation on vertical as well as horizontal planes.

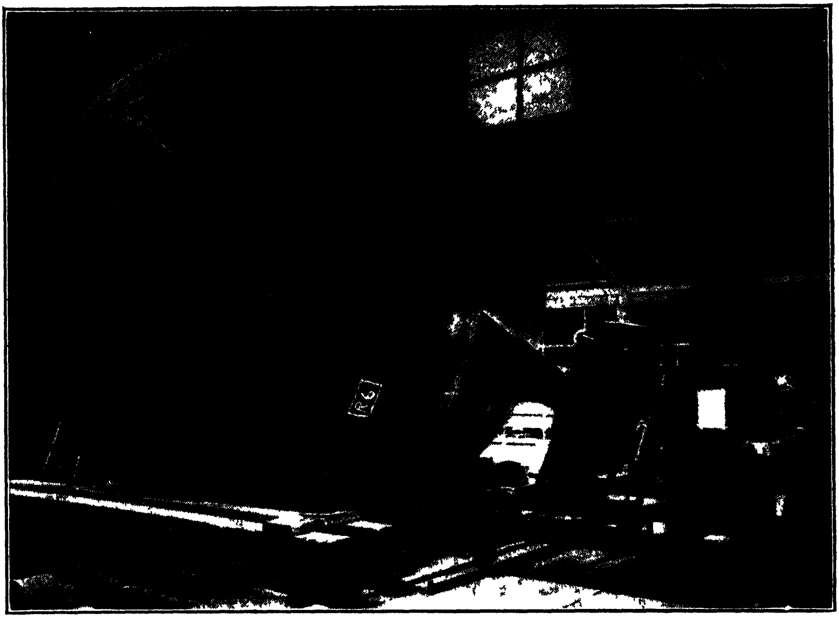


FIG 12 Truck-mounted, low-voltage X-ray unit used for making radiographs of a huge steel sphere used in dam construction (Courtesy General Electric X-Ray Corporation)

Radiography of steel $1\frac{1}{2}$ inches thick and castings of corresponding size require X-ray tube voltages of the order of 200,000 volts. For heavier and thicker sections, 400,000 volts have been used. The X-ray machines using these higher voltages are often portable and may be mounted on cranes or trucks. A truck-mounted industrial unit for inspecting heavy welds and for castings is illustrated in Fig. 12.

During World War II two new developments were made that have greatly extended the horizon on the application of X radiation. These developments are the 1- and 2-million-volt X-ray machines and the betatron. These devices have increased the hardness and power of penetration of X radiation and have reduced the required time of exposure for radiography.

The Million-Volt X-Ray Tubes. The new million-volt X-ray industrial units involve a new principle of action, a new type of tube, a new type of transformer, and a new form of assembly. The relative size of the new X-ray tubes and their construction are illustrated in Fig. 13.

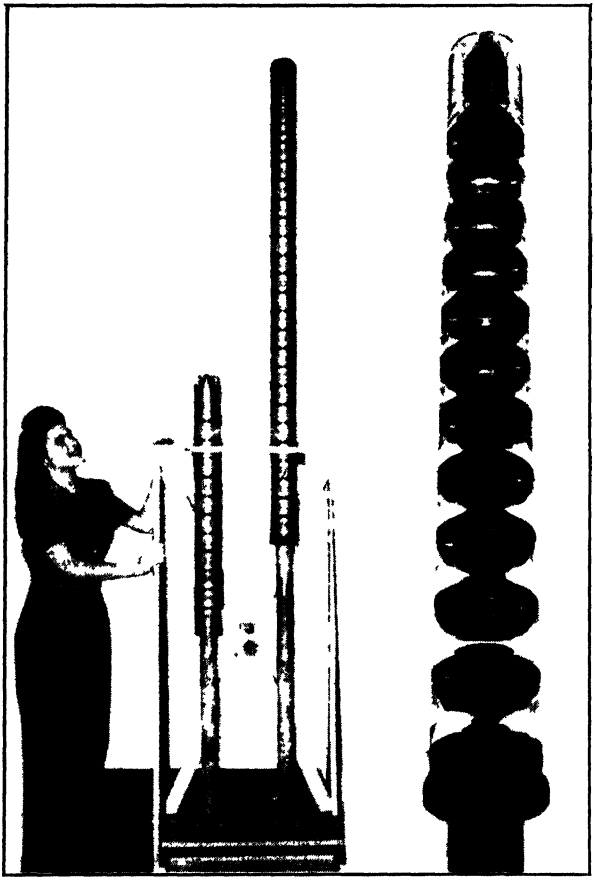


FIG. 13. Construction and comparison of the size of 1- and 2-million-volt X-ray tubes. (Courtesy General Electric X-Ray Corporation.)

These tubes have a filamentary tungsten cathode, a copper-backed tungsten target at the lower end of the extension chamber, and cylindrical accelerating electrodes in each of the 12 or 24 intermediate sections. The accelerating electrodes serve to distribute the high potential along the length of the tube. The inside walls of the tube are sand blasted to eliminate dangerous field current which might result from the application of increased voltage to each tube section. The assembly

of the unit is shown in Fig. 14. The long X-ray tube is placed at the center of a cylindrical resonance transformer. A cylindrical steel shell encloses the transformer and the main section of the X-ray tube. The space inside the steel shell is filled with freon gas at a pressure of 60 pounds per square inch for electrical insulation. The long anode end of the X-ray tube projects from the lower end of the case making it available for insertion into hollow cavities (circular and other shapes) for emitting the high-frequency X radiation. The circuit for the million-volt X-ray unit is given in Fig. 15. Power is supplied to the resonance transformer by a frequency changer set consisting of a 60-cycle synchronous motor driving a 180-cycle generator. The use of the resonance-type transformer reduces the weight because of the elimination of the iron core, the elimination of insulation space between the core and the high-voltage winding, and simpler tube connections.

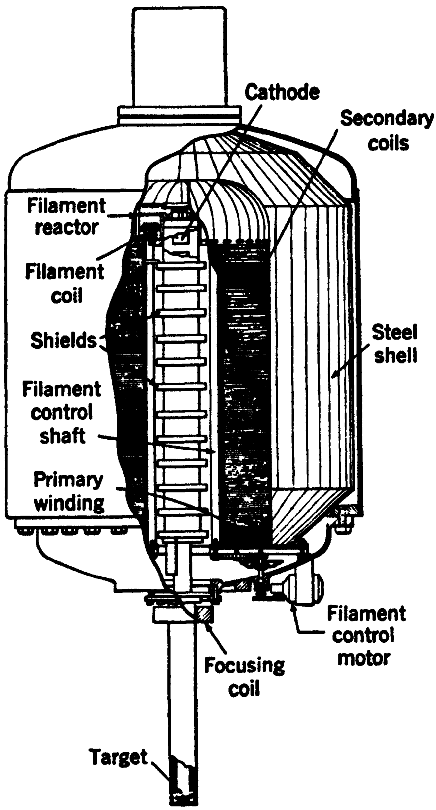


FIG. 14. Sectional drawing of a multisection X-ray tube concentrically mounted within a high-voltage resonance transformer. The unit has a diameter of 3 feet and a total length of 6 feet. (Courtesy General Electric X-Ray Corporation.)

The entire million-volt unit weighs 1800 pounds and is only 6 feet long and 3 feet in diameter. The 2-million-volt unit weighs 5100 pounds and is 8 feet long (11 feet with tube extensions) and 5 feet in diameter. The transformer has a low-voltage winding consisting of two flat coils of rectangular wire and a high-voltage coil consisting of 243 thin flat sections (for the 2-million-volt unit) spaced apart for cooling. A nearly uniform potential gradient is secured by spacing the upper coils more closely than those at the lower end. The focusing coil near the bottom of the main cylindrical

case employs the principle of magnetic focusing to bring the electrons to a focus on the target.

Power is applied to the X-ray unit by controlling the field of the 180-cycle generator. The operator raises the generator field gradu-

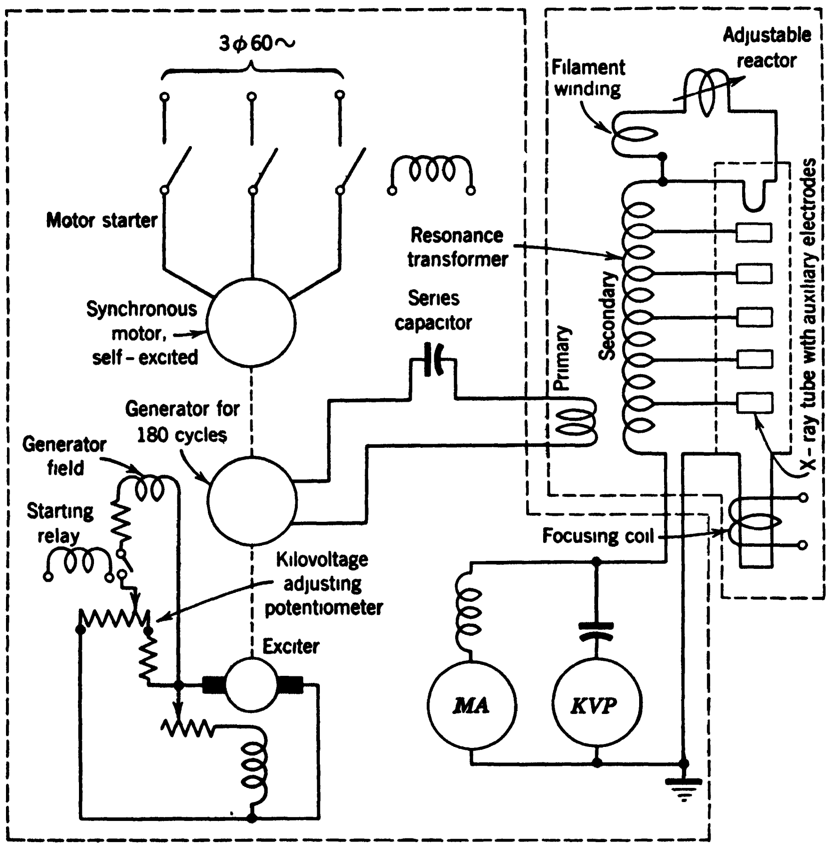


FIG. 15. Simplified wiring diagram for the X-ray unit of Fig. 14. The section in the upper right corner is contained in the tank. The rest of the equipment is a relay-rack-mounted control system. (Courtesy General Electric X-Ray Corporation.)

ally and continuously until the desired operating potential is reached. A water system for cooling the target of the tube and the freon gas is an integral combination of two radiators, a fan, two motors, a water pump, and protective relays. The electrical system of these units is so thoroughly interlocked that it is impossible to produce X radiation unless all factors are in order.

The method of supporting and applying the million-volt X-ray units is illustrated in Figs. 16 and 17. The latter figure shows how shells and bombs placed inside a cylindrical rotating turntable were moved in and out of the X-ray chamber for radiography. The 1-million-volt unit is suitable for radiographing up to an 8-inch thickness of steel and the 2-million-volt unit can be used for steel sections up to 12 inches thick. As to relative speeds, for a focal film distance of 4 feet on an 8-inch steel casting, the 2-million-volt unit is 100 times faster than the 1-million-volt machine. A picture can be taken through steel 12 inches thick in 2 hours at a distance of 3 feet with the 2-million-volt unit. An interesting radiograph taken with the 2-million-volt unit is shown in Fig. 18. A comparison of the effectiveness of various voltages in X-ray work is shown in Fig. 19.

FIG. 16. One-million-volt X-ray unit being used to inspect a large casting. (Courtesy General Electric X-Ray Corporation.)

Adequate protection must be given to all human beings who work near X-ray equipment. This protection must cover the hazards due to (1) high voltage and (2) the harmful effects of X radiation. The most effective bar to X radiation is sheet lead and it is often used to surround all parts of the X-ray tube except a window for the radiation. It is generally used for doors into X-ray exposure chambers and sometimes it is used to surround the entire room where X rays are used. A special form of plaster for the walls of the X-ray chamber is often used because of its lower cost. Also thick concrete walls are fairly effective for screening X rays. The advice of experts in the X-radiation field should always be secured before installation of X-ray equipment.

The Betatron. The betatron is an induction electron accelerator wherein electrons are accelerated in a magnetic field instead of by

direct application of a high potential. The basic idea of accelerating electrons by magnetic induction was patented by J. Slepian in 1927. In the years following this announcement Wideroe, Walton, and Steenbeck developed equations giving the necessary conditions for an electron accelerator but were not able to produce a machine that

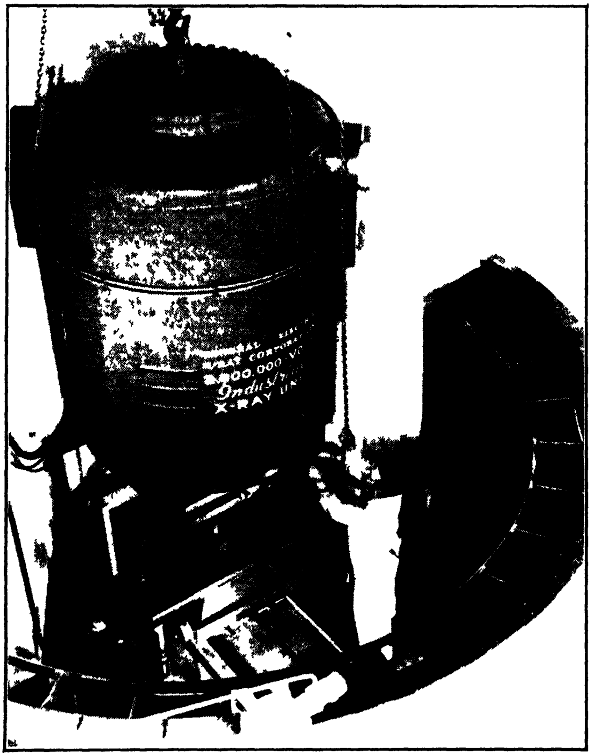


FIG. 17. Two-million-volt industrial X-ray unit used to inspect bombs and shells (Courtesy General Electric X-Ray Corporation)

would work. It remained for Dr. D. W. Kerst of the Physics Department of the University of Illinois to invent a satisfactory working model of the betatron (announced in 1940). Kerst assisted in the design and development of a 20-million-electron-volt (mev) commercial betatron in the early forties. A larger 100-mev commercial betatron was completed in 1945.

The basic principle of the betatron is relatively simple. In part *a* of Fig. 20 a circular iron core is surrounded by a single turn of wire *t*. If a changing flux is made to pass through the iron an emf will be

induced in the turn of wire and, if the wire circuit is closed, a circulating electron current will result. Since a current is a movement of electrons, it is obvious that, if electrons are released in the path shown by t while the flux changes in the core, these electrons will be induced to travel around the core in a curved path and will be accelerated as long as the flux is changing in the same direction. One problem in applying this simple principle is to make the electrons follow a con-

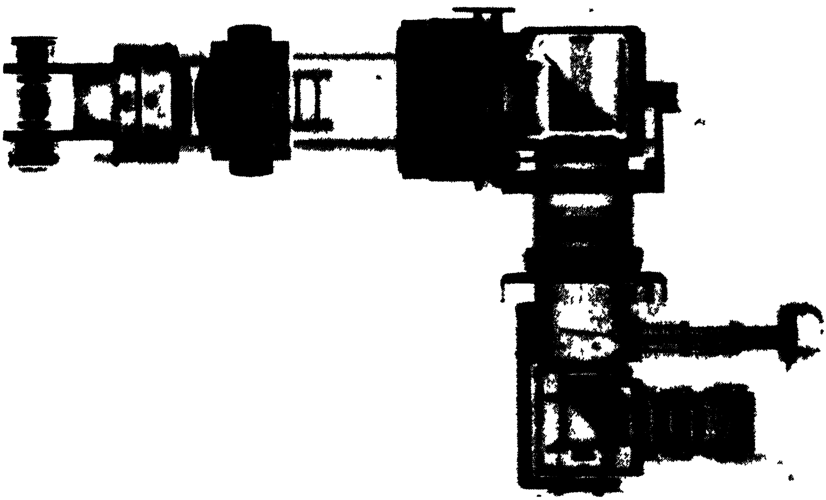


FIG. 18. X-ray of a German military periscope taken with a 2000-kilovolt X-ray unit showing extreme sensitivity and wide ratings to different densities. (Courtesy General Electric X-Ray Corporation.)

stant circular path. The changing flux for the betatron is produced in a three-legged transformer as shown in b of Fig. 20. A "donut"-shaped accelerating tube is placed around the central leg of the transformer in the plane of de. Exciting coils are placed on the top and bottom sections of the central leg. The transformer is excited by a resonance type of circuit as indicated in part c of Fig. 20. The resonant circuit on the transformer may be inductively coupled as shown, or a capacitive type of coupler may be employed.

Some of the details of the operation of the betatron are indicated in Fig. 21. The heart of the device is a torus or "donut" of glass which surrounds the changing stream of magnetic flux. The cathode and target are combined into a single unit. A heated tungsten cathode supplies electrons for an electrostatic gun which fires a stream of elec-

trons tangentially into the central circular path. After arriving in this path, the changing magnetic induction accelerates the flying electron in a circular path. When the desired velocity of the electron has been acquired, a transient current passed through a one-turn coil near

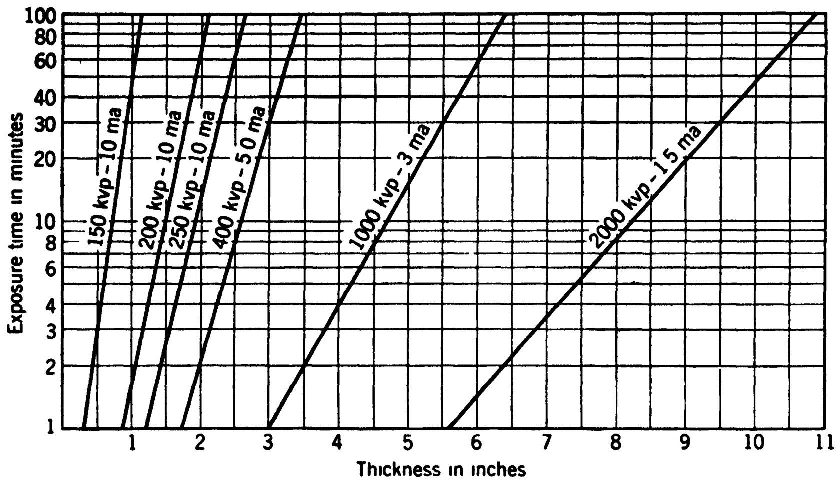


FIG. 19. Comparison of penetrating power of X radiation under different accelerating potentials. (Courtesy General Electric X-Ray Corporation.)

each pole face having a diameter slightly less than the optimum orbit serves to force the electron into a circular path of larger diameter (beam expanded) so that it hits the target and produces X rays. The flux change in the transformer follows the sine wave shown at the right

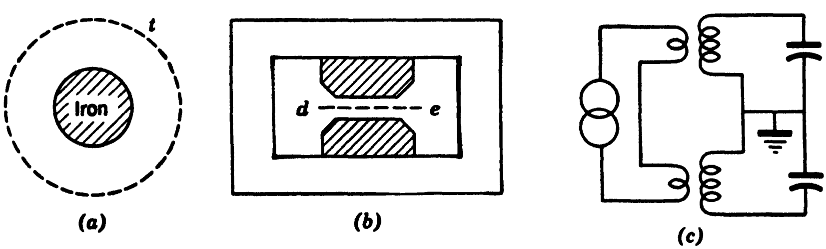


FIG. 20. Schematic construction and circuit of a betatron.

of Fig. 21. In operation the electron beam is injected at point A (near zero) on the sine curve and the beam is deflected or expanded at any point later (up to B), depending on the desired energy in million-electron-volts.

The problem of maintaining the accelerating electrons in an optimum circular path until they hit the target has been suggested. It is apparent that as the electron stream whirls and accelerates in a curved path the momentum of the electrons increases and they tend to move in a spiral path which would soon carry them outside the "donut." An electron moving at right angles to a magnetic field is acted upon by a constant force which causes it to move in a curved path (circle for constant field H). In the present problem the force of the stray field

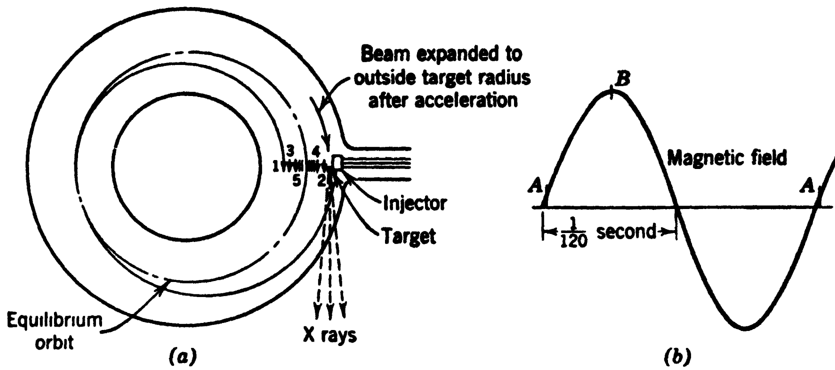


FIG. 21. (*a*) Schematic construction of the X-ray torus in a betatron. (*b*) Sine-wave magnetic field.

in which the electron stream moves tends to hold the path circular but, since the electron velocity and momentum is constantly increasing, the circular path will be a spiral (increasing diameter) if the field H is constant. Therefore, in order to hold the diameter of the electron path constant, it is necessary to have H increase at the same rate as the momentum of the electrons. This necessary condition has been secured by the proper shape of the iron pole faces and magnetic disks in the air gap giving a stray field flux of the necessary value along the optimum circular path of the electrons.

A 20-mev commercial betatron is illustrated in Fig. 22. The design of the betatron makes possible acceleration of electrons to very high energies with low-voltage equipment. In the 20-mev unit electrons are given the same energy they would receive if they were accelerated by a potential difference of 20 million volts, while the voltage required to create the necessary field for their acceleration by magnetic induction is about 100 volts per turn on the magnet coil. The maximum intensity of the 20-mev X rays is in the direction of the electron stream as it strikes the target. This intensity falls off 50 per cent at 4½

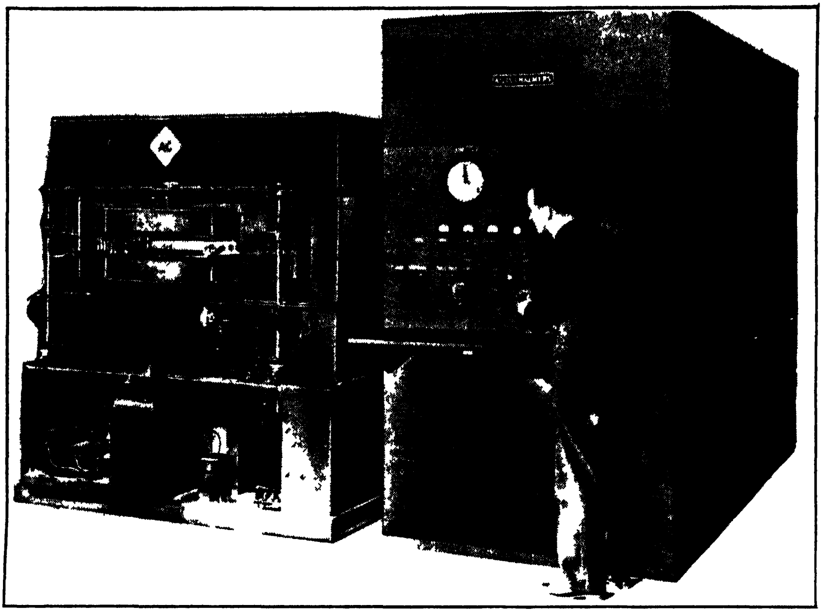


Fig. 22 20-mev betatron and control unit (Courtesy Allis-Chalmers Manufacturing Company)

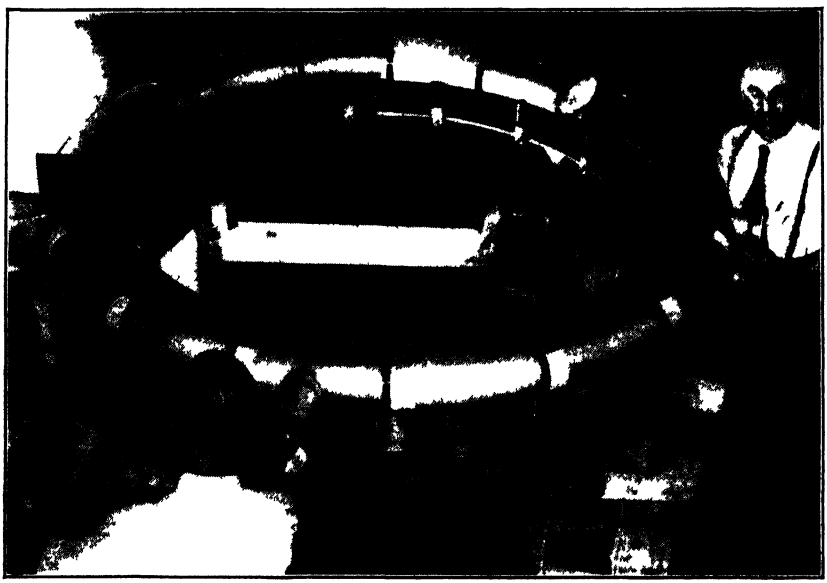


Fig. 23. View of the tube for a 100-mev betatron. (Courtesy General Electric X-Ray Corporation)

degrees from the center and somewhat more gradually at larger angles. The focal spot of the betatron is very small, about 0.05 inch high by less than 0.005 inch wide. This small focal spot accounts for a sharpness in detail of radiographic work when the betatron is used.

A 100-mev betatron was completed by the General Electric Company in 1945. This unit weighs 130 tons and is about 9 feet high, 6 feet wide, and 15 feet long, and it uses a 24,000-kilovolt-ampere capacitor for power factor correction. The X-ray tube is made of 16 molded glass sections connected end to end. The transformer magnet is operated on 60-cycle current with a flux density at the orbit of 4000 gauss. The electrons are injected with a voltage ranging from 30 to 70 kilovolts and, if allowed to remain in the 66-inch-diameter circular orbit for the entire quarter-cycle, they circle the magnetic flux 250,000 times, acquiring on each revolution an average additional energy of about 400 electron-volts. Thus in a period of $\frac{1}{240}$ of a second each electron travels over 800 miles. By means of the pulsing field they may be removed from the circular orbit at any time during the quarter-cycle, delivering energy varying from 1 mev up to the full 100 mev. The X-ray beam output has been as high as 2600 r (roentgens) at 100 mev, dropping to about 7 per cent of this value at 20 mev. A view of the tube for a 100-mev betatron is shown in Fig. 23.

In addition to its application in the radiographic field, the betatron holds much promise for use in both nuclear research and medical therapy. The reader will find detailed explanations of the theory, construction, and operation in the references cited at the end of this chapter.

PROBLEMS

1. Explain the operation of the Villard circuit shown in Fig. 4.

2. An X-ray therapy tube operates continuously at 250,000 volts direct current with a current of 10 milliamperes. If 90 per cent of the energy must be conducted away by water cooling, what will be the rate of water flow for a temperature rise (water) of 40 degrees F?

3. If 1,000,000 volts is applied between the cathode and target of an X-ray tube, what would be the velocity of the impinging electron calculated by formula 9 of Chapter I? How does this compare with the velocity of light? Explain.

4. Assume that the "donut" tube in a betatron has a mean diameter of 1 meter and that an electron in the tube gains energy at the average rate of 200 electron-volts for each revolution. What will be its velocity starting from rest after 0.001 second?

REFERENCES

SLEPIAN, J., United States Patent 1,645,304, October 11, 1927.

KERST, D. W., *Phys. Rev.*, **58**, 841 (1940).

KERST, D. W., and R. SERBER, *Phys. Rev.*, **60**, 53 (1941).

GIRARD, J. P., and G. D. ADAMS, *Elec. Eng.*, May 1946.

WESTENDORP, W. F., and E. E. CHARLTON, *J. Applied Phys.*, **16**:10, 581–593 (October 1945).

Chapter XXII

SPECIAL PHOTO APPLICATIONS

Oscilloscope.* A cathode-ray oscilloscope is an instrument designed for the analysis of electrical circuits by a study of the wave forms of voltage and currents at various points. The instrument may be employed to study any variable, within the limits of its frequency response characteristic, that can be converted into electrical potentials. Common variables are sound, vibration, light, and all forms of variable impedances such as resistance, inductance, and capacitance.

The basic elements of an oscilloscope are a cathode-ray tube, amplifiers, and a power supply. The construction and general theory of the cathode-ray tube was covered at the end of Chapter IV and the reader may wish to review that discussion before proceeding with this chapter. The frequency limitation of an oscilloscope is determined by the electron transit time across the face of the deflection plates in the cathode-ray tube and by the functioning of the amplifier circuits. In the cathode-ray tube the transit time is of the order of 0.001 microsecond so that little error will be present in frequencies up to 100 megacycles. Amplifier circuits do have frequency limitations and they become a controlling factor in applying the oscilloscope to many problems.

The basic theory of action of the cathode-ray tube as applied to the oscilloscope may be reviewed by observing the right side of Fig. 1. The electron beam moving perpendicularly to the page passes between the parallel deflection plates HH' and VV' and forms a luminous spot on hitting the fluorescent screen. With zero applied potentials to the deflection plates and the proper adjustment of the tube circuits the luminous spot should fall at the center of the screen. If a varying potential (a-c) is placed across plates HH', the luminous spot will be caused to travel back and forth along the horizontal line hh'. If the frequency of the variation of voltage is 16 cycles or more, the hori-

* The terms oscilloscope and oscillograph are used interchangeably for this device. The contraction "scope" is frequently used.

zontal trace will appear as a solid motionless line because of the persistance of vision of the human eye. In a similar manner, if a variation of voltage is applied only across deflection plates V and V', the moving luminous spot will create a vertical line or trace as indicated by vv'. When varying potentials are placed across both pairs of deflection plates simultaneously, a picture is created on the screen which gives information regarding the wave form of the applied signals. In general, some independent or known signal is applied across the plates

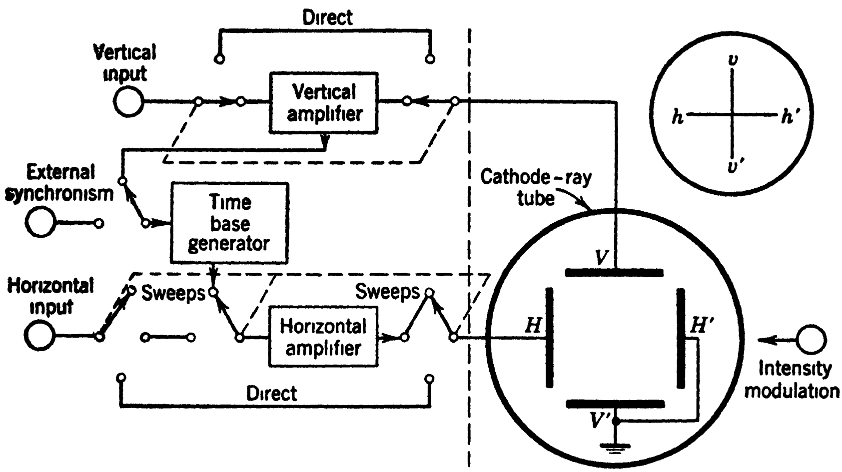


FIG. 1. Block diagram of a simple oscilloscope.

HH' while an unknown or dependent signal is applied across the plates VV'. The former is called the horizontal signal and the latter the vertical signal. For the determination of wave shapes a synchronized sawtooth signal is used for the horizontal while for the study of frequency and of phase angle a sine-wave a-c signal is used.

The block diagram of a simple oscilloscope is shown in the complete Fig. 1. The circuit is simplified and the cost of the device has been reduced by grounding H' and V'. The vertical input signal is fed between the post so marked and the ground. If this signal is of sufficient strength it is supplied directly to the deflection plates V and V', but if the magnitude is of insufficient value to produce a satisfactory deflection it may be passed through an amplifier for controlling the final signal strength applied to the plates. In a similar manner an external horizontal signal may be applied through the lower circuit group to plates direct (lower position of ganged switches) or via an amplifier

(middle switch position). For wave analysis it is generally desired to synchronize the "sweep" signal with the applied vertical input signal. This is accomplished by controlling the time base (sawtooth) generator from the vertical input with the connection as shown in Fig. 1. The latter may be considered as an internal horizontal input. For some applications it is desirable to control or modulate the intensity of the luminous spot on the screen by some external signal. This re-

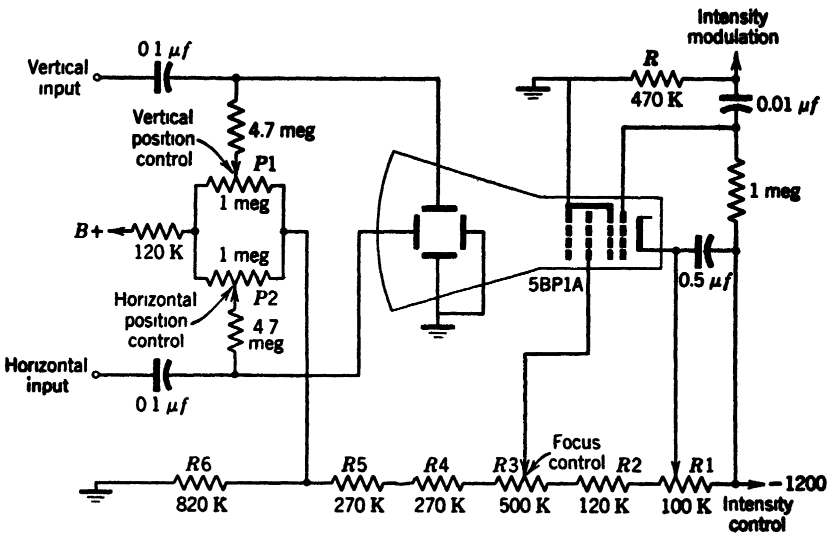


FIG. 2. Simplified schematic of cathode-ray tube circuits.

sult is accomplished by the intensity modulation indicated on the right of diagram. Intensity modulation in this circuit is effected by feeding a signal into the control grid-cathode circuit of the cathode-ray tube. A better understanding of the oscilloscope may be obtained by expanding the block diagram of Fig. 1 into component schematic circuits.

A cathode-ray tube circuit for a commercial oscilloscope, showing the magnitude of the components, is given in Fig. 2. The second anode and the deflection plates are held at or near ground potential by ground connections. The cathode and control grid are maintained at a potential of approximately −1200 volts, giving a maximum electron-accelerating potential of 1200 volts, from potential divider resistors $R1$ through $R6$. Focusing control is attained by connecting the second anode to the variable resistor $R3$. The luminous-spot intensity control is attained by varying the voltage between the cathode and the

control grid across a section of resistor $R1$. The position of the focal spot may be centered between the pairs of deflection plates by adjustment of potentiometers $P1$ and $P2$. It will be noted that the potential between the ends of the potentiometer resistances varies from negative (point between $R5$ and $R6$) to B+ (above ground).

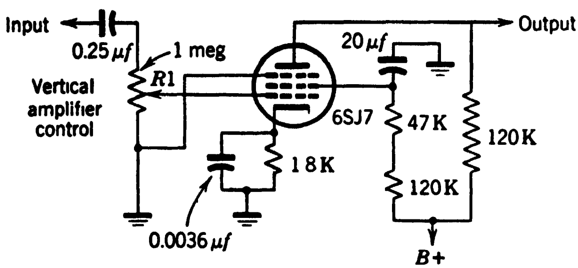


FIG. 3. Simplified circuit of an amplifier for an oscilloscope.

The amplifier circuits for the vertical and horizontal inputs to the deflection plates use pentodes as indicated in Fig. 3. The time base or sawtooth generator uses the circuit and control as indicated in Fig. 4. This is a relaxation oscillator circuit using a thyratron. The voltage supplied to cathode and anode circuit is determined by the voltage

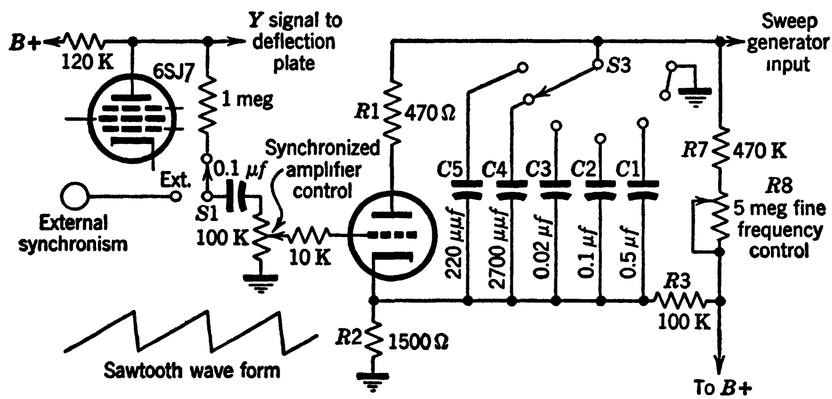


FIG. 4. Simplified schematic of a sweep circuit.

drop across the potential divider $R2$ and $R3$. For the setting shown, capacitor $C4$ charges through resistors $R8$ and $R7$ until the thyratron reaches its firing potential. At this point, $C4$ discharges rapidly, lowering the potential across the cathode-anode circuit below the conduction point. Then $C4$ recharges slowly and the process repeats, giving the

sawtooth wave form of voltage as shown in the lower left-hand corner of Fig. 4. The instant of firing (trigger action of grid) may be controlled by the connection to vertical amplifier (left) and thus give synchronous timing with the vertical input signal. Again synchronous timing from an external signal may be attained by movement of switch *S*1 to the left. An excessive synchronizing signal impairs the wave form of the sweep oscillator. Coarse frequency control can be ob-

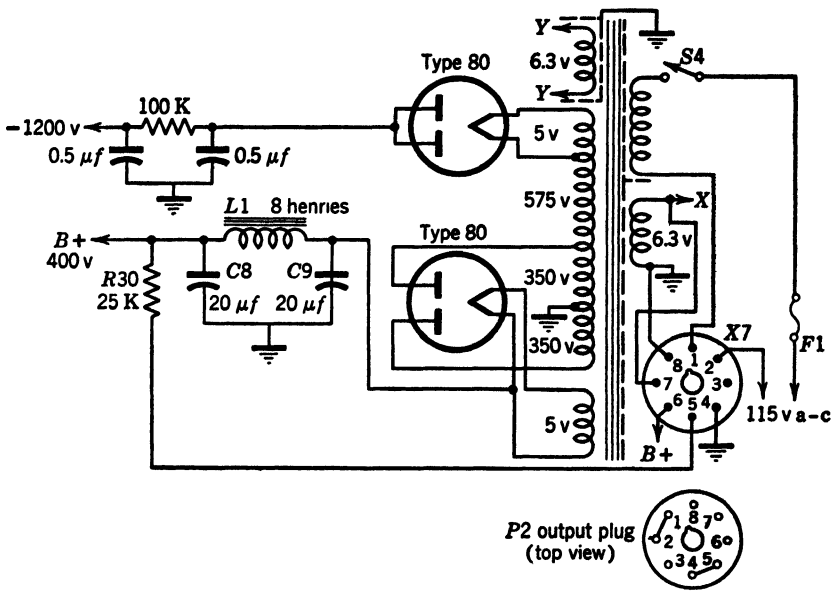


FIG. 5. Simplified circuit of the power supply for an oscilloscope.

tained by the capacitor selection of switch *S*3 and fine adjustment of frequency control can be effected by variation of resistor *R*8.

The schematic circuit for the power supply of the oscilloscope is given in Fig. 5. The top tube giving half-wave rectification furnishes the high-voltage (1200 volts) supply through a resistor and capacitor filter for the electron gun circuit. The lower tube gives full-wave rectification for a 400-volt supply for the tubes and deflection plates in the instrument plus a reserve for operating some electronic equipment accessory to the oscilloscope. The output plug in the lower right corner is for a take-off of B supply for accessory equipment and serves as a safety "disconnect" of the 115-volt, a-c supply. The safety feature results because this plug must be pulled out to remove the case from the instrument.

A commercial oscilloscope (oscillograph) which embodies the preceding circuit is illustrated in Fig. 6

The action of the oscilloscope when studying a wave form plotted

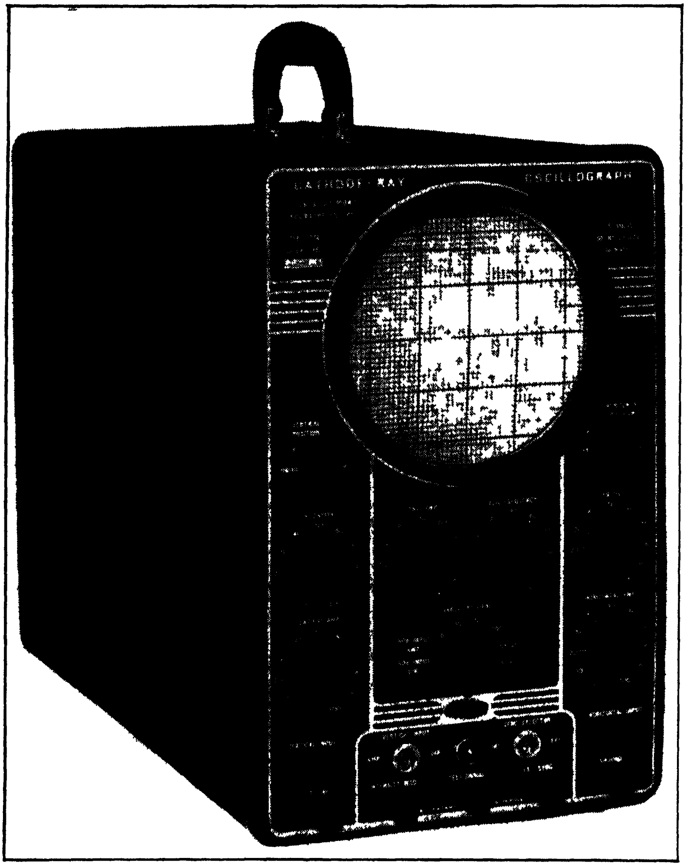


FIG 6 A commercial oscillograph (Courtesy Allen B Dumont Laboratories, Inc)

against time is illustrated in Fig. 7 A sine wave of voltage is applied to the vertical input (left) and a sawtooth wave of the same frequency is applied across the horizontal input. The magnitude of the vertical and horizontal sweep is governed by respective input signal control. For comparative and quantitative measurement of the signals a transparent screen containing cross-section lines is placed on the face of the cathode-ray tube.

When the signal inputs to both the horizontal and the vertical deflection plates are a-c voltages, the resulting action and pattern on the screen are illustrated in Fig. 8. This form of pattern is known as a Lissajous figure, named after the nineteenth-century French scientist. The Lissajous pattern may be traced by joining intersections from like numbered points on the signal. Lissajous figures are used for deter-

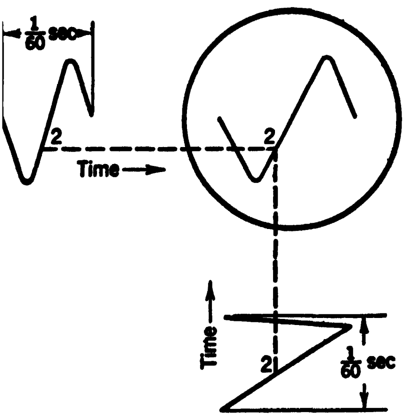


FIG. 7. Projection drawing of a sine wave applied to a vertical axis and a sawtooth wave of the same frequency applied simultaneously on a horizontal axis.

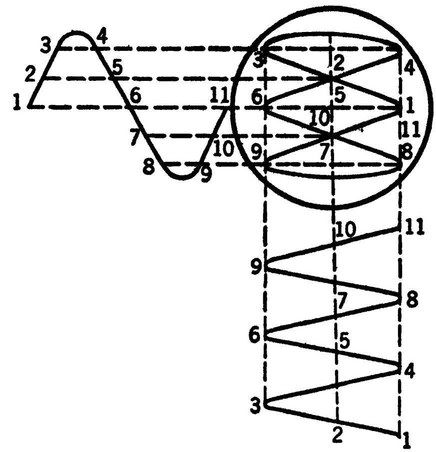


FIG. 8. Projection drawing showing the resultant Lissajous pattern when a sine wave applied to the horizontal axis is three times the frequency applied to the vertical axis.

mining the frequency of unknown signals. Such frequency determination requires the use of one signal of known frequency which is impressed on the horizontal input. If the signals of the same frequency and magnitude and 90 degrees out of phase are impressed across the vertical and horizontal inputs, the trace will be a circle, as shown in part 1 of Fig. 9. Thus the circle becomes the pattern for a frequency ratio of 1/1. A horizontal frequency ratio of 3/1 is illustrated in Fig. 8 and several other ratios in the views of Fig. 9. The frequency relationship is determined by the ratio of the number of loops touching two mutually perpendicular sides such as AB and BC of part 5 of Fig. 9. The algebraic rule for the determination is

$$\frac{\text{Frequency on horizontal axis}}{\text{Frequency on vertical axis}} = \frac{\text{Number of loops intersecting } AB}{\text{Number of loops intersecting } BC}$$

The phase-angle difference between two signals of the same frequency can be determined by the use of Lissajous figures produced by impressing these signals on the horizontal and vertical input to the oscilloscope. In making this determination it is important (1) that the luminous spot be centered on the screen of the cathode-ray tube, (2) that both horizontal and vertical amplifiers have been adjusted to give exactly the same gain, and (3) that the calibrated scale be set to coincide with the displacement of the signal along the vertical axis. The calculation for the angle of phase shift is made by use of the formula

$$\sin \theta = \frac{Y \text{ intercept}}{Y \text{ maximum}}$$

where the Y intercept is the magnitude of the $+Y$ intercept of the trace on the Y axis and the Y maximum is the peak or maximum $+$ value attained by the trace. Several views of patterns and the calculation of the phase-angle difference are illustrated in Fig. 10.

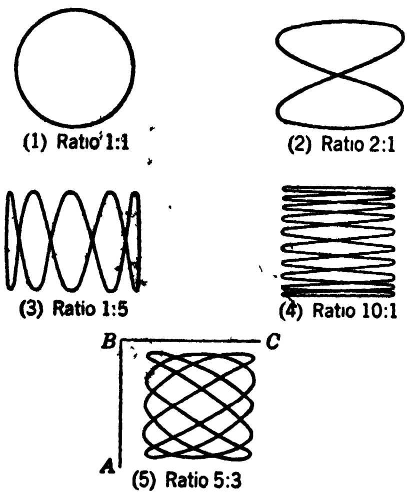


FIG. 9. Lissajous patterns.

The frequency of a series of pulses and the duration of a single

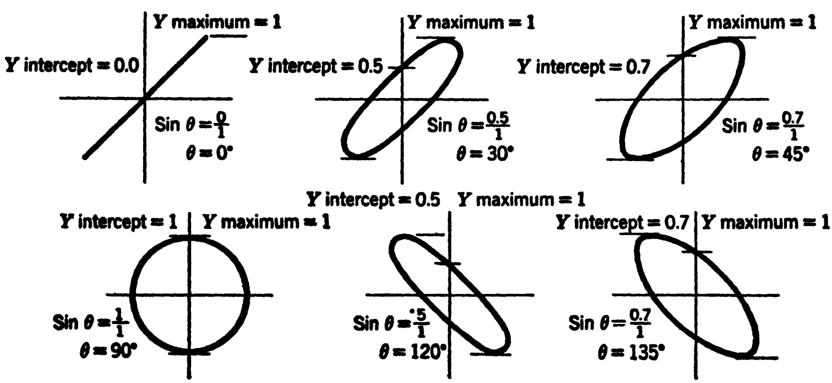


FIG. 10. Examples showing the use of the formula for determination of phase difference.

pulse may be determined by impressing the impulse signal upon a known sine wave through intensity modulation. Intensity modulation is the result of applying a signal of varying potential to the control grid of the cathode-ray tube, thereby varying the intensity of the trace at the frequency of the signal applied. The result of such intensity modulation is illustrated in the oscillogram of Fig. 11.

Electron Microscope. The electron microscope is a modified and enlarged cathode-ray tube for magnifying minute objects. The optical form of microscope is limited in its effective resolving power by the wavelength of visible light to about 1000 diameters. With ultraviolet light the effective resolving power can be increased to 1500 diameters. Electrons have wave properties which permit them to be used to give resolving power fifty times as great as for light rays.

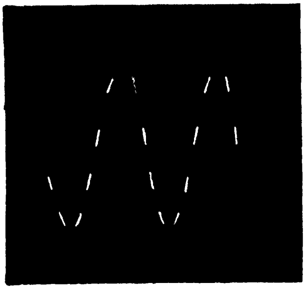


FIG. 11. Example of intensity modulation with a sine-wave signal. (Courtesy Allen B. Dumont Laboratories, Inc.)

The first American commercial model of an electron microscope was designed and built by Zworykin and Morton in 1940. The principle of operation of this new device in comparison with an optical microscope is illustrated in Fig. 12. In the light microscope (*a* of the figure), light rays from a lamp are formed into a parallel beam and directed on a specimen S by the condenser lens L_1. The image of the specimen then falls on the objective lens L_2 which focuses and magnifies it, producing an enlarged image I_1. Part of this enlarged image is further magnified by the projector lens L_3. The twice-enlarged image I_2 is seen by the eye.

In the electron microscope (*b* of Fig. 12), the source of action is a hot cathode which emits electrons. Beneath the cathode there is an electron gun and anode which gives the electrons a high velocity downward. A coil or solenoid L_1 produces a magnetic field which has a focusing action and bends the paths of these electrons into a parallel beam directed on the specimen S. The electrons in the beam are affected in a varying degree, depending on the density of different parts of the specimen. Those that pass through are brought to a focus by the field of a second coil L_2 and form an enlarged image I_1. The electron rays that form a section of this image are in turn deflected and magnified by the field of a third coil L_3 and caused to form a larger image I_2. Since this image I_2 is formed by an electron beam it is not

visible. Accordingly, a fluorescent screen is placed so that the electron beam falling on it produces a visible image. If a directly viewable image is not desired, a photographic film is used in place of the fluorescent screen.

An improved commercial model of an electron microscope is shown in Fig. 13. This model is 75 inches high and 24 inches wide. It is capable of making magnifications varying from 100 up to 20,000 diameters via direct viewing. In many cases the picture (negatives) may be given a useful photographic enlargement up to 100,000 diameters. Greater enlargement does not furnish any additional or useful information. A picture taken with the electron microscope showing a total magnification of 94,000 diameters is shown in Fig. 14. The location of several of the operating units of the electron microscope is indicated on Fig. 13. The vacuum in the microscope column is created by an oil-diffusion pump. Pumping down to an operating vacuum of approximately 10^{-6} millimeter Hg requires from $1\frac{1}{2}$ to 2 minutes. The specimen is placed in the vacuum chamber near the top of the column and the cover to this opening is held closed by the vacuum. Since electrons will not penetrate glass the specimen holder consists of a film of nitrocellulose 1/1,000,000 centimeter thick held on a fine wire mesh. The photographic film is inserted into the vacuum column by a gate within easy reach of the operator. The microscope column, the oil-diffusion pump, and the power supply are built into a single cabinet as shown in Fig. 13. The voltage regulator and the mechanical first-stage pump are separate and can be located in another room.

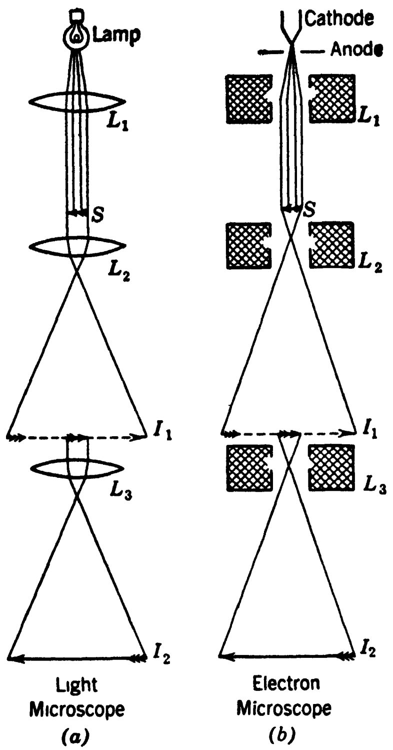


FIG. 12. Comparison of an optical and electron-microscope lens system.

The magnetic flux lenses in the electron microscope consist of iron-encased coils. The iron casing for the coils has an air gap on the inside

so shaped as to give flux paths at the center of the coil which will deflect the electron in the desired directions. A suggestion of the construction is given in Fig. 15. The magnification secured in the electron

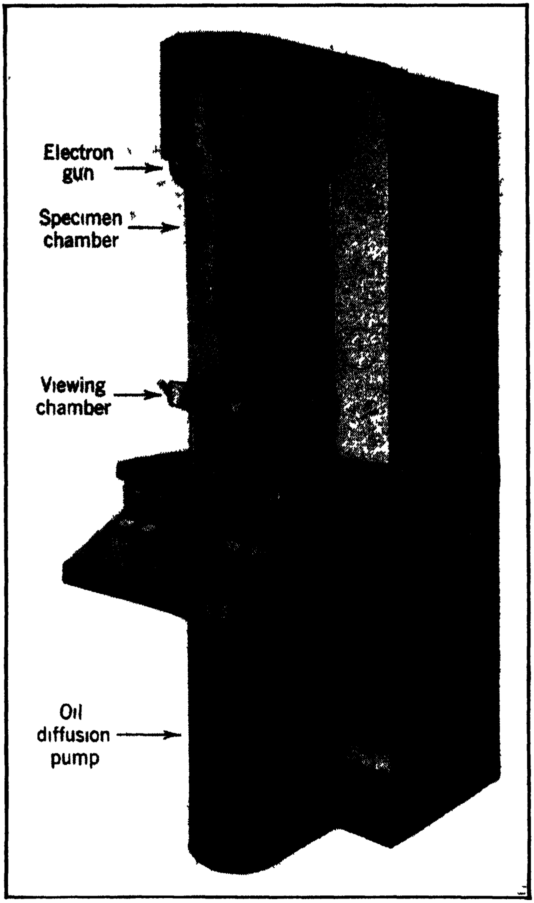


FIG. 13. Universal-type electron microscope. (Courtesy Radio Corporation of America.)

microscope can be controlled by variation of the current in the magnetic lens coils and by variation of the velocity created by the electron gun.

The electron microscope can be used as an electron diffraction camera. This application requires the removal of the specimen from the upper chamber and its insertion in the lower chamber which is in

the position of L_3 in Fig. 12. The diffraction pattern is photographed in the same manner and position as for magnified images.

The introduction of the electron microscope has opened a large field in scientific and industrial research. It has made possible the study of the structure of bacteria and virus not discernible by the optical microscope and the study of minute particles of matter in the fields of chemistry, metallurgy, botany, and food products.

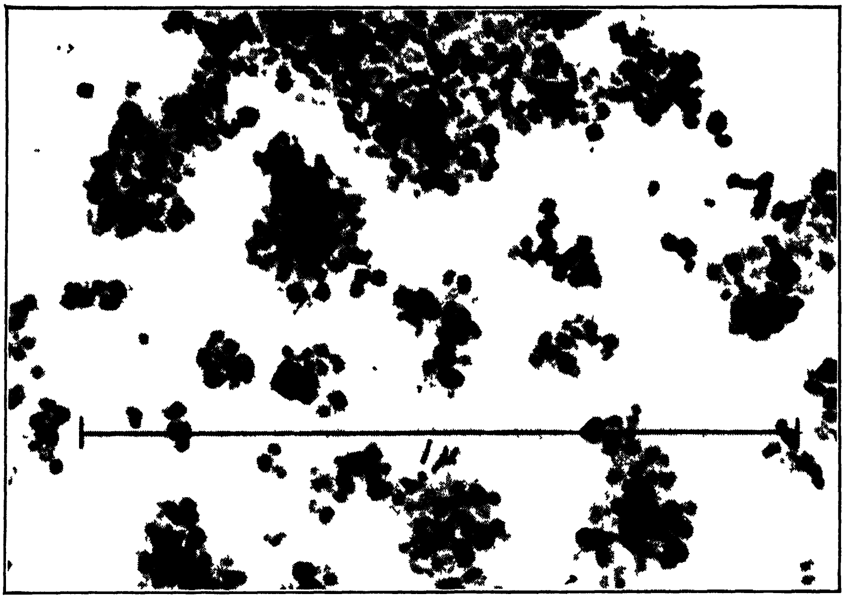


FIG. 14. Colloidal titania (titanium dioxide) photographed with the electron microscope and enlarged to 94,000 diameters (Courtesy Radio Corporation of America.)

Stroboscope. The stroboscope is an instrument for studying or observing a periodic or varying motion by means of light periodically interrupted. Such an instrument comprises a luminous-arc discharge tube containing gas or vapor which is caused to flash by periodic transient currents. If a rotating shaft, wheel, or gear is given a flash of light once for each revolution it will appear to be at rest, or if the flashes occur at $1/n$th of its revolutions it will appear likewise to be stationary. Again, if the period of the flashes is slightly more than the period of the revolutions, the device will appear to be turning forward slowly at a speed determined by the difference in time of the flashes and period of one revolution. Similarly, a flash period less

than a revolution period will give the illusion of a slow backward motion.

This phenomenon of the stroboscope has useful applications, chief among which is the measurement of speed. When the rotating machine appears at rest the rate of flashing is the same or a submultiple of the speed, and if the flashing rate is adjustable and calibrated the stroboscope becomes a tachometer. A commercial device known as a Strobotac built for speed measurement is shown in Fig. 16. This instrument has a calibrated dial and gives an accuracy of ± 1 per cent of the dial for readings above 900 revolutions per minute when the Strobotac is standardized in terms of a frequency-controlled power line. The advantages of this type of speed measurement are that (1) no power is absorbed from the mechanism and (2) it can be used to measure the speed of machine elements inaccessible to ordinary tachometers. Thus speeds may be determined for small electric motors and delicate mechanisms which would slow down or stop if minute amounts of power were drawn from them. Likewise, the speeds of inaccessible gears or wheels in a complicated mechanism may be obtained. A second important application of the stroboscope is the slow-motion study of high-speed motion for determining vibration, tension, chattering, whip, and other irregularities that may be present in rotating and repetitive forms of motion. Thus mechanical defects and troubles may be located and remedied in many kinds of industrial machines.

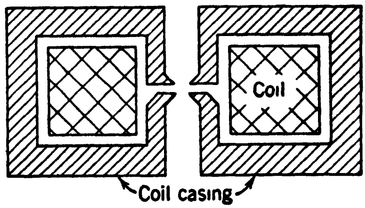


FIG. 15. Cross section of a magnetic focusing lens.

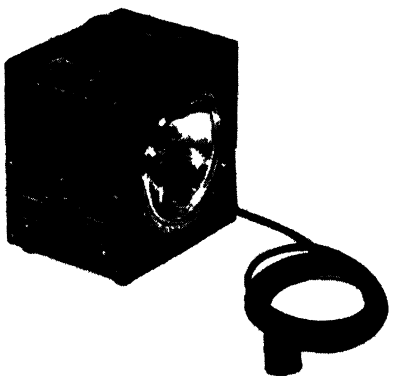


FIG. 16. A commercial form of stroboscope called a Strobotac. (Courtesy General Radio Company.)

Several different tubes and circuits may be employed in designing a stroboscope. The particular tube or circuit chosen will depend upon the amount of light desired, the period of the flash, the application, the cost, and other factors. Theoretically, any cold-cathode gaseous or vapor tube may be employed.

For low brilliance of light a small neon glow lamp (see page 108) may suffice or the firing of a thyratron gives some light. Long-column gaseous and mercury-pool and mercury-vapor tubes are often employed. The most widely used tube for stroboscopic applications is the strobotron which was shown in Fig. 35 in Chapter VI. This tube was developed at The Massachusetts Institute of Technology by Edgerton and Germeshausen. The special features of this tube are that it operates at a relatively low voltage, conducts a very high peak current of short duration, giving a very brilliant flash, and is easily controlled through two grid electrodes.

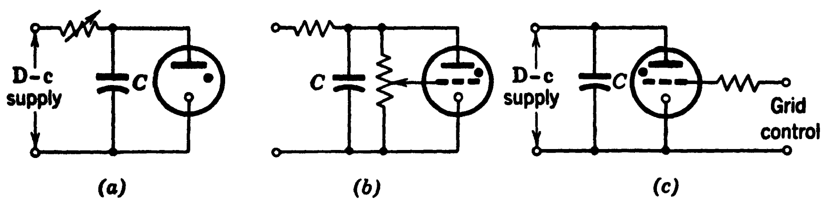


FIG. 17. Control circuits for flash bulbs.

Three simple types of circuits for the operation of a stroboscope are suggested in Fig. 17. In all these and other circuits a large capacitor C is needed to supply a heavy discharge current and to give a brilliant and quick flash. The circuit of a is a relaxation oscillator circuit using a two-electrode gaseous tube. The flashing period may be adjusted by the variable resistor which controls the time for the capacitor to charge up to the breakdown potential of the tube. Part b of Fig. 17 suggests the use of a self-exciting grid control for initiating the breakdown of the flash tube. All present-day applications use a flash tube with an internal or external grid for initiating the firing because of the relatively high breakdown potential between the cold cathode and anode. The circuit of part c, Fig. 17, suggests some form of external grid control. Such grid control may be produced by a mechanical contactor located on the machine which is being studied by stroboscopic light. In other cases the grid control may be effected by an auxiliary oscillator or timing circuit.

High-Speed Photography. High-speed photography is the art of taking pictures of high-speed motion by using brilliant flashes of light of a few microseconds' duration. This new art has been made possible by investigations and developments of Edgerton and Germeshausen. High-speed photography utilizes the general principles discussed under the stroboscope but requires more brilliant flashes of shorter time

duration. One commercial form of a flash lamp for high-speed photography is illustrated in Fig. 18. A long flash tube or column is made in a helical form in order to concentrate the light in a smaller area. The helical tube has a small diameter which gives a quick deionization time and short flash. A typical circuit for use in connection with this tube is given in Fig. 19. This circuit embodies a principle frequently employed in stroboscopes for initiating the discharge of the flash tube. This principle utilizes an ignition type of transformer for impressing a high transient voltage across the grid. In the lower right-hand corner of Fig. 19 a capacitor $C3$ is charged to the potential across resistor $R2$. The closing of switch $S1$ permits $C3$ to discharge through the primary of transformer $T1$ which, in turn, induces a high voltage (approximately 10,000 volts) between the grid and the cold cathode of the flash tube. This transient initiates cold-cathode emission and capacitor $C2$ discharges through the flash tube. Flash tubes are now employed in conjunction with portable auxiliary equipment by the press and commercial photographers for on-the-spot photography in the place of photoflash bulbs. A commercial device of this type is illustrated in Fig. 20.

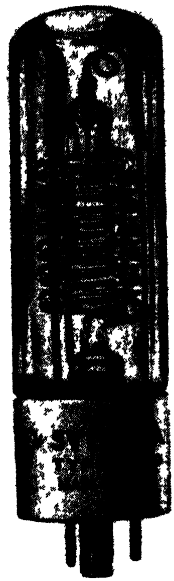


FIG. 18. Flash tube for high-speed photography. Typical operating voltage, 2250 volts; flashing rate, 6 per minute; light output peak, 12,000,000 lumens. (Courtesy Sylvania Electric Products, Inc.)

High-speed photography is employed in research to study the conditions taking place in explosions, destruction of equipment by projectiles, projectiles in flight, and other transient phenomena. In some cases a rapid series of flashes is used to show changes and to determine the velocity of projectiles and other moving objects. An excellent photograph of this type is shown in Fig. 21.

Iconoscope. The iconoscope is a special form of cathode-ray tube used in television transmission for "picking up" a scene and converting it to an electrical signal. The principal parts of an iconoscope are the mosaic, signal plate, collector, and electron gun. The position of these parts in the tube is illustrated in Fig. 22. The mosaic consists of a large number of small photosensitive globules deposited on one face of a thin sheet of insulation. The globules are spaced a very small dis-

tance apart on the sheet so as to be insulated from each other. On the opposite face of the insulating sheet is a conductive film, the signal plate. Because the insulating sheet is thin, there is considerable

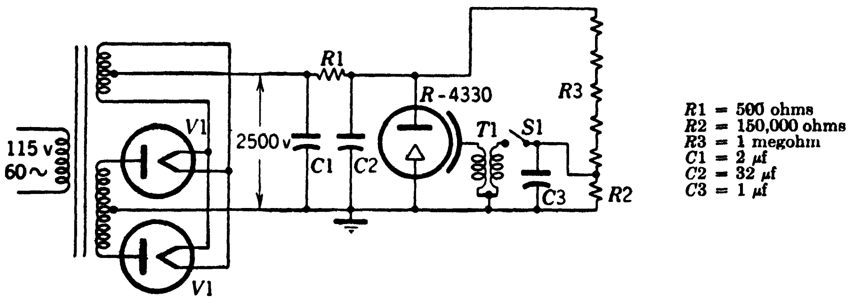


FIG. 19. Circuit for a photoflash operation.

capacitance between the globules and the signal plate. The collector is a conductive coating on the inner surface of the tube walls which collects electrons emitted by the mosaic.

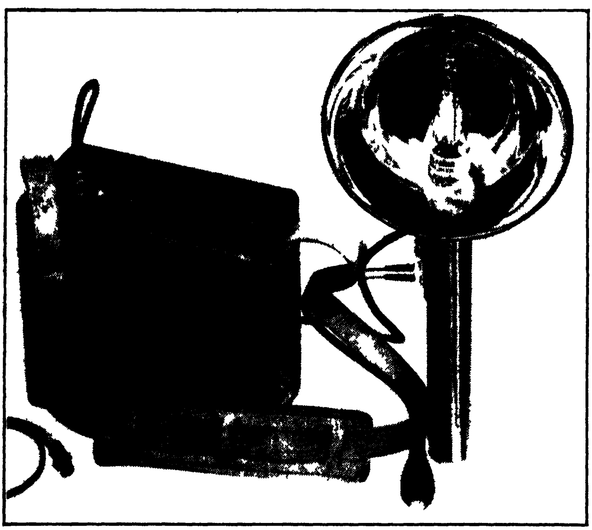


FIG. 20. Portable photoflash equipment for photography. (Courtesy Sylvania Electric Products, Inc.)

In operation of the iconoscope, an image of a scene is focused by a lens system on the mosaic, and the beam of electrons provided by the electron gun is made to scan the image. As the beam moves over the

image, there is generated at the signal plate a voltage whose magnitude at any instant depends on the image brightness at the point where the beam is at that instant. This voltage can be used as the video signal for television transmission of the scene viewed by the iconoscope. The process by which the iconoscope generates this voltage can be described briefly as follows.

FIG. 21. Multiflash photograph showing the pattern made by a golf club as it moves through the stroke. (Courtesy H. E. Edgerton.)

Consider first the action of the tube when the mosaic is scanned by the beam with no illumination on the mosaic. When the electron beam strikes a globule, the globule emits secondary electrons, the number of secondaries being several times larger than the number of beam electrons striking the globule. Some of these secondaries return almost immediately to the globule, the rest escape and go either to the collector or to other parts of the mosaic. Because the globule is insulated, its potential will change in the positive direction if the number of electrons escaping from it is greater than the number of electrons

flowing to it. The number of electrons that escape depends on the potential of the globule, the number being less, of course, the more positive the globule is. Hence, if the beam is on the globule a sufficiently long time, the globule will be driven to a positive potential at which the number of electrons escaping is equal to the number of electrons arriving. In the usual operation of the iconoscope, the time required for the beam to pass over a globule is long enough for the globule to attain this potential. The value of this potential for typical operating conditions is about 3 volts positive with respect to the collector.

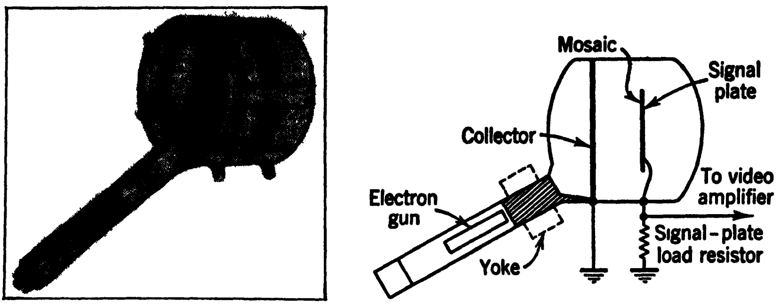


FIG. 22. Iconoscope. (Courtesy Radio Corporation of America.)

After the beam passes the globule, some of the secondary electrons emitted from the rest of the mosaic fall on the globule. The arrival of these electrons changes the globule potential in the negative direction to a new value. In a typical operating condition, this value is about 1½ volts negative with respect to the collector. With no light on the globule, the globule stays at this negative potential until the next time the beam strikes, when the globule again releases electrons and rises to its maximum positive potential of approximately +3 volts.

Consider now the action of the tube when the mosaic is scanned with part of it illuminated. During the time between contacts with the beam, both an illuminated globule and an unilluminated one receive electrons from the rest of the mosaic. Both globules, therefore, charge in the negative direction during this time. The illuminated globule, however, at the same time emits electrons, the emission being caused by the light on this globule. The illuminated globule, therefore, does not fall to as negative a potential as the unilluminated one does. Hence the next time the beam strikes, the illuminated globule does not have so far to rise to reach +3 volts. As a result, less charge is released to the collector when the beam strikes the illuminated globule than

when the beam strikes the unilluminated one. The difference in charge is approximately proportional to the difference in illumination.

Now consider the action of the tube when an image is projected on the mosaic and scanned by the beam. As the beam moves over the mosaic, varying amounts of charge flow from the mosaic to the collector, the amount of charge flowing at any instant being a measure of the light on the globules where the beam is at that instant. In other words, a video signal current flows between the mosaic and the collector. It can be seen that, since the beam current to the mosaic is constant, the video signal current must complete its circuit path through the signal-plate load resistor and the capacitance between signal plate and mosaic. The voltage developed across the load resistor by this signal current is the video signal output of the iconoscope.

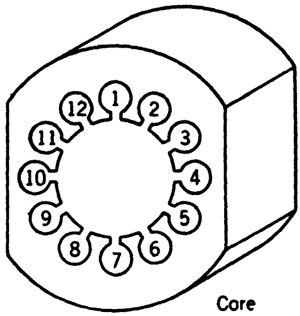


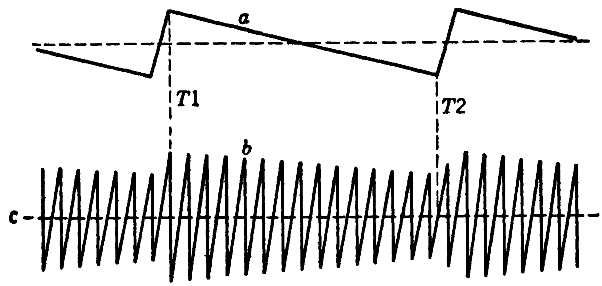
FIG. 23. Magnetic core for a yoke.

It also can be seen that, when the beam moves from a dark portion of the image to a brighter portion, the electron current from the mosaic to the collector decreases. The output voltage, therefore, changes in the negative direction. Hence the signal output of the iconoscope is of negative polarity: a highlight in the image is represented by a negative value of signal voltage; a shadow in the image is represented by a positive value.

The focusing and deflection of the electron beam is produced by magnetic fields. The focusing is secured by a solenoid coil which surrounds the neck of the iconoscope. Two pairs of coils for producing the magnetic deflection are wound in the slots of an iron core (Fig. 23). The focusing and deflection coils and their core are assembled into a single unit called the yoke which slips over the neck of the iconoscope as shown in Fig. 22.

The horizontal and vertical deflection of the electron beam is produced by sawtooth signals having wave forms as indicated in Fig. 24. The horizontal sawtooth signal causes the beam to move rapidly from left to right across the image on the mosaic and then return in near zero time to the left. While the horizontal sweep signal is causing the beam to scan the image in the horizontal direction, the vertical sawtooth signal of a lower frequency causes the beam to move slowly from the top to the bottom of the mosaic and then return with a nearly instantaneous sweep to the top. The result of these combined sweep

signals is to cause the beam of electrons to scan a rectangular area of the image focused on the mosaic as indicated in Fig. 25. This entire scanning process takes place in a period of approximately $\frac{1}{30}$ second. The ratio of the frequency of the horizontal to the vertical signal is about 500/1, which means that 500 horizontal lines or sweeps take place in $\frac{1}{30}$ of a second. During any one of these horizontal sweeps



a = vertical frequency modulating voltage
b = voltage across condenser $C3$
c = average value of voltage across $C3$

Fig. 24. Sweep signals for television.

the light intensity viewed on the mosaic may change from light to dark in $\frac{1}{500}$ of the distance along the sweep. This means that the period of a maximum light change may be

$$\tfrac{1}{500} \times \tfrac{1}{500} \times \tfrac{1}{30} = 1.33 \times 10^{-7} \text{ second}$$

which corresponds to a frequency of 3,750,000 cycles. Thus the *video* or image signal picked up in the iconoscope is of the order of several megacycles.

Television. Television is the art of transmitting electrically the image of an object in a fraction of a second. Such transmission may be via wire (coaxial cable) or radio. The essential elements of the television process are some form of electronic camera such as an iconoscope and a cathode receiving tube. Successful transmission of pictures via television requires (1) perfect synchronizing of the sawtooth sweep signals in the iconoscope and in the cathode-ray tube receiver and (2) the distortionless transmission of the video signal picked in the camera to the intensity control grid of the cathode-ray tube. If at every instant of time the electron beam in the receiving tube falls upon a spot on its fluorescent screen corresponding to the spot being hit by the electron beam in the transmitting camera and if the intensity of the

receiver beam corresponds to the intensity of the signal being picked up at that instant, the received image should correspond to the image in the transmitting camera.

Television has many commercial, industrial, and military applications in addition to its natural field in entertainment. Television receivers on several floors of a department store may show a style show taking place in another part of that store. The factory manager may

FIG. 25. Picture scanning in television.

view processes taking place in various places in that factory. Dangerous processes in the manufacture of explosives may be viewed through television by operators located in safe surroundings at a distant point. Guided missiles and pilotless planes may be followed and controlled through the use of television.

REFERENCES

ZWORYKIN, MORTON, RAMBERG, HILLIER, and VANCE, *Electron Optics and the Electron Microscope,* John Wiley & Sons, 1945.

GERMESHAUSEN, K. J., and H. E. EDGERTON, *Elec. Eng.,* **55**, 709 (1936).

ZWORYKIN, V. K., "Iconoscopes and Kinescopes in Television," *RCA Review,* July 1936.

VACUUM ELECTRONIC TUBES *

Tube Names	Function	Application	Type of Control	Type of Cathode
Kenotron *Diode*	Rectifies low values of current (often at high voltage)	Rectification in radio, X ray, and electrostatic precipitation	None	Hot
X Ray	Produces X-ray radiation	Medical diagnosis and therapy Examination of industrial products	None	Hot
Pliotron May be triode, tetrode, or pentode	Amplification and high frequency oscillation	High-frequency heating Carrier current Control circuits Radio receivers and transmitters	Electrostatic Grid	Hot
Cathode Ray	Produces visual indications by controlled electron beam	Circuit analysis Television	Electrostatic or Electromagnetic	Hot
Magnetron	Ultra-high-frequency oscillation	Short-wave radio	Electromagnetic field	Hot
Klystron	Ultra-high-frequency oscillation	Used in military radio and detection devices	Resonance chamber	Hot
Vacuum Phototube	Used in applications where audio-frequency response is important	Sound movies	Light variation	Photoelectric

* This table and the one which follows condensed from *Electronics at Work*, Westinghouse Electric Corporation, 1943.

GAS-FILLED ELECTRONIC TUBES

Tube Names	Function	Application	Type of Control	Type of Cathode
Rectigon *Tungar*	Low voltage rectification	Battery charging	None	Hot
Phanotron Mercury-vapor rectifier	Rectification of moderate amounts of current at voltages up to 20,000	Radio transmitters and industrial applications	None	Hot
Thyratron May be gas triode or gas tetrode	Control and controlled rectification of moderate amounts of current at voltages up to 22,000	Welding control Timing circuits Motor control Voltage regulation	Electrostatic Grid	Hot
Glow Tube	Illumination, rectification, and regulation at low current, low voltage	Switchboard indicating lights Control circuit voltage regulators Protection of supervisory control circuits	None	Cold
Grid-Glow Tube Gas triode	Controls small amounts of power	Safety control of furnaces	Electrostatic Grid	Cold
Pool Tube Mercury-arc rectifier			None	Pool (usually mercury)
Grid-Pool Tube *Excitron* Mercury-arc rectifiers	Rectification, conversion, and control of large amounts of power	Power for aluminum and magnesium production Electrochemistry Transportation Resistance welding	Electrostatic Grid	Pool (usually mercury)
Ignitron			Resistance ignition	Pool (usually mercury)
Gas Phototube	Process and quality control Detection devices	Sound movies Pinhole detector Cutting, winding and printing control	Light variation	Photoelectric

INDEX

Ampere, definition of, 5
Amplidyne, general, 242
Amplidyne current control, 423
Amplidyne positioning control, 426
Amplidyne speed control, 428
Amplidyne voltage control, 422
Amplification, theory of, 76
Amplification factor, 65
Amplifiers, classification of, 209
  couplings for, 212
  push-pull, 211
  servo, 258
  vacuum-tube, 205
Amplitude control, thyratrons, 122
Anodes, construction of, 45
Antihunt circuits, 267
Arc discharge, 105
Arc-welding control, 419
Atom structure, 2, 3

Barrier-layer cells, 171–173
Battery-charger control, 415
Beam-power tube, 85, 86
Betatron, 443–449
Breakdown voltage, 102, 103

Capacitron, 147
Cathode, construction of, 26
  oxide-coated, 30, 117
  thoriated-tungsten, 27
  tungsten, 27
Cathode-ray tube, construction of, 86–89
Cathode sputtering, 112
Charge, definition of, 4
  on electron, 7
  image positive, 20
  negative, 4
  positive, 4
  space, 4
Child's equation, for triode, 68
  for vacuum diode, 48

Cold-cathode tubes, 148–151
Color selection, 400
Color sensitivity, 162
Conduction, electric, 5
  in gases, 101, 111
  in liquids, 6
  in solids, 5
Constant-voltage circuit, 234
Contact emf, 39
Contactors, electronic, 235
Control systems, closed cycle, 250
  simple, 249
Copper-oxide rectifier, 185–188
Counting devices, 393
Crystal rectifiers, diode, 184
  general, 180
  germanium, 181
  silicon, 181
Current control, 423

Deionization, 100
Dielectric heating, 318–323
Diode, application of high-vacuum, 54
  vacuum characteristics of, 46–49, 52, 53
Discharge, arc, 105
  glow, 104, 105
  Townsend, 101, 102
Dynamic characteristic, 64
Dynamic response, 167

Edison effect, 19
Electric current, definition of, 5
Electric-field intensity, 8
Electron, 2
  free, 5
Electron affinity, 19, 20
Electron ballistics, 6
Electron deflection, in combined field, 15
  in electric field, 10
  in magnetic field, 12, 13, 14

Electron emission, 22
  effect of gas on, 32
  high-field, 22
  secondary, 23
  thermionic, 23, 25
Electron focusing, 14, 17
Electron microscope, 22, 459
Electron-multiplier tubes, 177
Electron optics, 16
Electron-ray tube, 92
Electron velocity, 8
Electron-volt, 9
Electronics, definition of, 1
Electrostatic field, definition of, 7
  distribution of, in triode, 59
  uniform, 9
Elevator-leveling control, 420
Emitters, application of, 33
  comparison of, 32
  effect of gas on, 32
Excitron, 138

Filters, full-wave, single-phase, 287
  general, 285–290
  half-wave, single-phase, 286
  L-type, 288–290
  smoothing, 285
Fluoroscopy, 434
Free-wheeling circuit, 227

Gaseous conduction, 96
Gaseous lighting units, 107
Gaseous rectifier diodes, 114
Glow discharge, 104
Glow lamps, 107, 148
Grid, construction of, 55, 56
  free, 60
  theory of, 55
Grid-control ratio, 120
Grid-glow rectifier, 151
Grid-glow tube, 152–154
Grid leak, 61
Grid of thyratrons, 127

Heat control, 349
High-field emission, 22
High-frequency heating, 297–326

Iconoscope, 465–470
Ignitron, anode firing, 141
  general, 140–146
  network firing, 142
  vs. thyratron, 147
  water-cooled, 145
Image positive charge, 20
Induction heating, 298–318
Inverse peak voltage, 113
Inversion circuits, 291–295
Ion, negative, 4
  positive, 3
Ion production, 97
Ionization by collision, 96, 97
Ionization methods, 97
Ionizing collision, 100
Ionizing potential, 99

Light-control devices, 387–405
Lighting control, 391, 417
Long-tailed pair, 237

Mean length of free path, of electrons, 96
  of molecules, 95
Metallic rectifiers, copper-oxide, 185–188
  general, 183–194
  magnesium-copper sulphide, 190–191
  selenium, 189–190
Molecule, structure of, 3
Multiple-field exciters, 245
Mutual conductance, 67

Negative space charge, general, 43
  neutralization of, 110
Nonlinear resistors, 234

Oscillator, Colpitts circuit, 219
  crystal, 220
  electron-coupled, 219
  Hartley circuit, 218
  vacuum-tube, 216
Oscilloscope, 451–458
Oxide-coated cathode, 30, 31

Parameters of vacuum tubes, 68
Paschen's law, 104
Peaking transformer, 225

Pentode, characteristics of, 82, 83
construction of, 81
potential distribution of, 81
Phanotron, 117
Phase-shift circuits, 225–227
Phase-shift control, 122, 124, 125
Photoconductive devices, 174–176
Photoelectric control devices, 387–405
Photoelectric pyrometer, 402
Photoelectric recorder, 405
Photoemission, general, 160
laws of, 161
Photosensitive films, 162
Phototubes, gaseous, 166–169
vacuum, 162–166
Photovoltaic devices, 169–173
Plasma, 107
Plate, definition of, 45
Plate resistance, r_p, 67
Position-control devices, 240, 426
Potential distribution, in beam-power tube, 86
in diode, 44
in electric fields, 9
in hot gaseous tube, 111
in pentode, 81
in tetrode, 78
in triode, 57
Potential gradient, definition of, 8
Power pack, 51
Precipitation, Cottrell process of, 362–366
electrostatic, 362–370
Precipitron, 368–370
Protection devices, 398

Q factor of circuits, 199

Ratio of rectification, 274
Rectification, full-wave, 51
half-wave, 50
theory of, in vacuum, 40
Rectifiers, commutating, 268
comparison of, 194
contact, 270
copper-oxide, 185–188
electrolytic, 271
magnesium-copper sulphide, 190
mercury-vapor, 116
Rectifiers, multielectrode metal-tank, 135–138
multiphase, circuits of, 279–284
rotary, 271
selenium, 189
single-phase circuits, 272–279
summary of, 285
tungar, 116
vibrating, 268
Register controls, 399
Regulex, 246
Relay tube, 152
Resistance welding, divisions of, 330
general, 328–360
Resonance, parallel, 201–203
series, 200
Richardson's equation, 25
Ripple factor, 277
Rotary amplifiers, 242
Rototrol, 246

Saturable-core reactors, 221–223
Schottky effect, 35
Screen grid, 77
Secondary emission, general, 23
in tetrode, 79
Selenium rectifier, 188–190
Selsyn system, alternating-current, 241
direct-current, 240
Series impedance transformer, 224
Servomechanisms, 251–267
Skin effect on copper, 198
Sorting devices, 396
Space charge, negative, 43
Spectrum radiant energy, 158–159
Speed control, 428
Speed regulation, 413
Stroboscope, 462–464
Stroboscopic lamps, 149
Synchro system, *see* Selsyn system

Television, 470
Temperature-control devices, 239
Temperature regulation, 414
Tetrode, construction of, 76
potential distribution of, 78
Thermionic emission, curves of, 28, 29
equation of, 25
general, 23

Thermistors, 235
Thyratron, general, 118
  hydrogen, 132
  shield-grid, 128, 131
Thyratron amplitude control, 122
Thyratron phase-shift control, 122
Time-control devices, 231-233
Timers, electronic, 341
Townsend discharge, 101, 102
Transconductance, 67
Trigger tube, 122
Triode, characteristics of, 62–64
  construction of, 55, 56
  conventional circuit of, 60
  limitations of, 73
  potential distribution of, 57
  types of, 69
Tungar rectifier, 116
Tungsten, curve of emission of, 28, 29
  thoriated, 27
Tungsten cathode, 27

Unilateral conductivity, 39
Utility factor, 275

Variable-mu tube, 66
Vibrating rectifier, 269
Voltage-divider circuits, 229
Voltage regulation, by amplidynes, 422
  electronic, 407–412
Voltage-regulator tube, 149

Welding, energy storage, 356
  seam and pulsation, 353
Work function, definition of, 21
  table, 21
  theories of, 34

X radiation, 430
X-ray equipment, 431
X-ray million-volt tubes, 440
X-ray tube, 430